Citation: Hicks JL, Uchida TK, Seth A, Rajagopal A, Delp SL. Is my model good enough? Best practices for verification and validation of musculoskeletal models and simulations of movement. Journal of Biomechanical Engineering 137(2): 020905, 2015. doi: 10.1115/1.4029304.

Is my model good enough?

Best practices for verification and validation of musculoskeletal models and simulations of movement

Jennifer L. Hicks, Thomas K. Uchida, Ajay Seth, Apoorva Rajagopal, Scott L. Delp

Department of Bioengineering, Stanford University, California, USA

Submitted to the Journal of Biomechanical Engineering

December 2, 2014

Address correspondence to:
Jennifer L. Hicks
James H. Clark Center, Room S-342
318 Campus Drive
Stanford, CA 94305
jenhicks@stanford.edu

Abstract

Computational modeling and simulation of neuromusculoskeletal (NMS) systems enables researchers and clinicians to study the complex dynamics underlying human movement. NMS models use equations, derived from physical laws and biology to help solve challenging real-world problems, from designing prosthetics that maximize running speed to developing exoskeletal devices that enable walking after a stroke. NMS modeling and simulation has proliferated in the biomechanics research community over the past 25 years, but the lack of verification and validation standards remains a major barrier to wider adoption and impact. The goal of this paper is to establish practical guidelines for verification and validation of NMS models and simulations that researchers, clinicians, reviewers, and others can adopt to evaluate the accuracy and credibility of modeling studies. In particular, we review a general process for verification and validation applied to NMS models and simulations, including careful formulation of a research question and methods, traditional verification and validation steps, and documentation and sharing of results for use and testing by other researchers. Modeling the NMS system and simulating its motion involves methods to represent neural control, musculoskeletal geometry, muscle–tendon dynamics, contact forces, and multibody dynamics. For each of these components, we review modeling choices and software verification guidelines; discuss variability, errors, uncertainty, and sensitivity relationships; and provide recommendations for verification and validation by comparing to experimental data and testing robustness. We present a series of case studies to illustrate key principles. In closing, we discuss challenges the community must overcome to ensure that modeling and simulation are successfully used to solve the broad spectrum of problems that limit human mobility.

1 Introduction

Almost every complex engineering product, from bicycles to aircraft, is designed using modeling and simulation. Engineers are confident using modeling and simulation for the design of these systems because the underlying mathematical models of materials and system dynamics have been tested in many applications over the past several decades, and the available computational tools have been validated relative to precise measurements of analogous real-world systems. Thus, modeling and simulation of mechanical systems have great impact on product design and engineering.

Modeling and simulation of biological systems are just beginning to make an impact in healthcare. For example, subject-specific computational fluid dynamics models of the heart are being developed to help diagnose coronary artery disease [1]. Despite their potential impact, however, computational models of biomechanical systems have yet to be applied on a wide scale in healthcare, largely due to the complexity and variability of biological systems combined with the heightened challenge of validation.

Computational modeling and simulation of the human neuromusculoskeletal (NMS) system, one branch of biological modeling and simulation, shows great promise for improving the diagnosis and treatment of the many conditions that limit human mobility. For example, NMS models can reveal internal muscle forces and joint loads for a wide range of scenarios, from activities of daily living like walking to high-performance maneuvers like sprinting and jumping. Internal muscle and joint forces are difficult or impossible to measure experimentally, but understanding how muscle forces coordinate motion is

essential for applications like designing assistive devices, planning rehabilitative treatment, and understanding the fundamental principles of human locomotion.

NMS modeling has grown rapidly in the last 25 years. A search on Google Scholar for biomechanical or musculoskeletal modeling or simulation produced fewer than 200 papers in 1990, about 500 papers in 2000, and nearly 2000 papers in 2013 (Figure 1). Further, motion analyses performed by the hundreds of clinical and research gait labs around the world represent a type of modeling that is now standard in the diagnosis of movement disorders in many hospitals, and is integral to a wide range of biomechanical studies.

In spite of this growth, NMS modeling and simulation have yet to be applied widely in clinical practice or medical device design, in large part due to gaps in validating these models to ensure their accuracy and reliability. Evaluating the validity of models and simulations for answering a specific research or clinical question is the responsibility of all researchers and developers of NMS models or simulation tools. Verification and validation are essential for convincing modelers and non-modelers alike of the utility of simulation results and overcoming the barriers modelers often face in publishing papers, attracting funding, and translating results to the real world. Appropriate validation is needed for modeling studies to have a broad impact, and is good scientific and engineering practice, helping to prevent the proliferation of erroneous conclusions.

1.1 Definitions

Several terms must be defined before reviewing verification and validation best practices for NMS modeling and simulation.

Model: We define a model as a set of mathematical equations that describe a physical system, along with the computational expression of these equations. For our purposes, the physical system is the human or animal neural and/or muscular system acting on a rigid multibody skeletal structure, possibly in interaction with external devices and/or the environment (e.g., the ground).

Simulation: A simulation is the result of using a model to study a specific motion or biophysical event. A kinematic simulation involves analyzing the motion of a system without regard to the forces causing or produced by this motion. A dynamic simulation can be forward or inverse. In a forward dynamic simulation, the model's equations of motion are integrated forward in time to predict the motion resulting from the application of forces. In an inverse dynamic simulation, the motion of a model is used to predict the forces that caused that motion.

Verification: We follow the ASME definition, where verification is "the process of determining that a computational model accurately represents the underlying mathematical model and its solution," [2] or "are we solving the equations correctly?"

Validation: We also follow the ASME definition, where validation is "the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model," [2] or "are we solving the correct equations?"

Calibration: Calibration is the process of choosing model and simulation parameters that provide the best match to experimental or other reference data. Models and simulations must be carefully calibrated before verification and validation are performed. Data used to calibrate a model cannot be used to validate a model or simulation.

Error, Accuracy, Uncertainty, Sensitivity, Credibility: Error is the *difference* between a measured or estimated value of a parameter and its true value. Conversely, accuracy quantifies the *agreement* between measured or estimated values and their true values. Uncertainty is a potential source of error, which can arise from a gap in knowledge about the biological or physical system under study, or from inherent variability in the subject or phenomenon being measured. Sensitivity is a measure of the effect of a change in a particular variable on the simulation outputs of interest. We define credibility as the degree of trust placed in a particular model or simulation for answering a specific research question. A simulation is most credible when the outputs of interest are insensitive to variables with high uncertainty. Anderson and colleagues [3] provide an excellent review of these concepts.

1.2 Objective

Most biomechanics researchers are aware of the importance of validation, but the field lacks best practices for the challenging process of verifying and validating NMS models and simulations. Several papers [3–5] have laid the groundwork, identifying principles and considerations, but these papers stop short of providing specific guidelines for NMS modeling and simulation. The knowledge and practices of how to best validate a biomechanical model and verify modeling software used in past research studies have not been adequately synthesized. The goal of this paper, therefore, is to establish practical guidelines for verification and validation of NMS models and simulations that researchers, clinicians, reviewers, and others can adopt to evaluate the accuracy and credibility of modeling studies.

In the sections that follow, we provide an overview of the verification and validation process in NMS modeling and simulation based on practices that have evolved in our research group and a review of the literature. We then detail best practices for verification and validation of each common component of a NMS modeling and simulation framework, from multibody dynamics to neural control modeling. We discuss 1) recommended modeling choices (e.g., appropriate level of modeling complexity) and verification and validation best practices for a given model or simulation output of interest; 2) how close a match is expected for comparisons between different types of simulation data and the corresponding experimental data; and 3) which variables have the greatest error, uncertainty, or influence and must be most rigorously tested in the validation process.

We illustrate the best practices through case studies. While these studies use the OpenSim [6] software platform for modeling and simulation, our guidelines are broadly applicable across the NMS modeling and simulation field, regardless of the software used. We close the paper with a discussion of key challenges for the field.

2 Verification and Validation Process

We break the verification and validation process for modeling and simulation into seven stages:

1) Formulate a research question that a model and simulation can answer;
2) Prototype your methods and create a verification and validation plan;
3) Verify your software;
4) Validate your results by comparing your model and simulation to independent experiments and other models;
5) Test the robustness of the study by evaluating the sensitivity of your results to model parameters and other modeling choices;
6) Document and share your model and simulation; and
7) Generate predictions and hypotheses that can be tested in the real world.

The verification and validation process begins with the definition of a suitable research question, continues through the process of designing and conducting your study, and extends beyond study completion, since sharing and documenting your results allows others to reproduce, extend, and test your models and simulations (Figure 2). Although we have enumerated a sequence of stages, the verification and validation process typically

requires iteration. For example, the validation plan should be complete before experimental data are collected, but the discovery of a particularly sensitive variable might necessitate additional experiments. We have geared this paper toward researchers performing modeling and simulation; however, the process we describe could be adapted to other types of modeling (e.g., finite element modeling of the musculoskeletal system) or to modeling and simulation for engineering or clinical applications. This section gives an overview of the entire process; more details for each potential component of a modeling and simulation pipeline are provided in Section 3.

2.1 Formulate a research question that a model and simulation can answer

Your research question must be well posed for modeling and simulation; otherwise, adequate validation will not be possible. Given the challenging nature of validating models and simulations, you must first assess whether answering the proposed research question will make a novel and important contribution to the field. This assessment is subjective, but several considerations can help establish importance and novelty, including 1) whether your research will improve our fundamental understanding of normal or pathological human movement; 2) whether your research will improve the diagnosis, treatment, or prevention of pathology or injury that inhibits human movement; 3) whether your research will enhance mobility, or performance; and 4) whether others will be able to reproduce, apply, and extend your work.

Next, you must determine whether modeling and simulation are necessary to answer your research question. In some cases, for example, an experimental analysis may be more appropriate, or sufficient experimental data may already be available to test your hypothesis. Conversely, you must also determine whether it is possible to test your hypothesis with a model or simulation. For example, the variables of interest must be robust outputs of the model or simulation so that you can draw credible conclusions in the face of model uncertainty. Further, there must exist a modeling and simulation framework capable of answering your research question, or you must have the expertise and resources to build one. In Figure 3, we introduce our verification and validation case studies [7–12] and describe how we formulated high-impact research questions that could be addressed with modeling and simulation.

2.2 Prototype your methods and create a verification and validation plan

The next step is to design your methods, including the modeling and simulation framework you will use and any experimental data you will collect, and create a verification and validation plan to ensure confidence in the analysis of your results and conclusions drawn (credibility). You should be able to answer "yes" to each of these questions before collecting data or generating models and simulations:

1. Do you understand how your modeling framework maps to the physical system of interest and how the framework is implemented?
2. Does your research question lie within the scope of intended uses of the model and the simulation framework?

3. Have you eliminated model complexity not required to answer your research question?
4. Have you identified modeling assumptions and their implications?
5. Have you collected or obtained to use as inputs to calibrate your model and generate simulations, or will you collect experimental data? Do you know the variability, error, and uncertainty in these measurements?
6. Do you have independent data to assess the variability, errors, and uncertainty identified above?
7. Can you address remaining variability, error, and uncertainty with sensitivity testing?

In Section 3, we review the common components of a modeling and simulation pipeline and for each component, provide information about the questions above, including making modeling choices to fit your intended use, common modeling assumptions and limitations, typical inputs along with their errors or uncertainties, and best practices for independent validation and sensitivity testing.

2.3 Verify your software

Before generating any results or conducting validation, you must verify your software to ensure that the computational model and underlying algorithms used to simulate the physiological or physical phenomenon are implemented correctly (i.e., the results match known standards). One cannot overstate the difficulty and importance of this step: if an algorithm has been implemented incorrectly, all resulting simulations will be incorrect and any conclusions drawn from these data will be contaminated. Paradoxically, it is not the

most substantial errors that are most problematic, since the erroneous results they produce are likely to be immediately apparent. Rather, it is the small errors that can be most pernicious, as they can produce results that are plausible and can, therefore, easily go unnoticed.

If the software has been designed in a modular fashion, the software engineer can perform verification for each component on its own (called "unit testing"), and can design higher-level tests for functionalities that depend on aggregates of components. A comprehensive verification suite ensures that inputs are bounded, only the state and unknowns can change, and computations—preferably for benchmark problems with known solutions—are repeatable as the code is updated. If modeling a physical system, the verification suite should test for physical principles such as energy conservation, including targeted tests where errors and miscalculations are likely.

The most effective technique for software verification is to employ existing software modules that have been independently verified and widely used for a range of applications, whether commercial or open source. If you choose open source modules, you and others can review and verify the code directly. There are well-established examples of widely applied and well-validated open source tools. Some that we have used successfully include LAPACK, GCVSpline [13], and IPOpt [14]. The OpenSim [6,15] and Simbody [16] software packages are also open source to promote community review and contributions. Publicly available modeling and simulation platforms should include a suite of verification tests for their modeling libraries and simulation algorithms. Even if you are not conducting the

verification yourself and are instead relying on existing testing suites, you should be aware of the key requirements for verification.

As mentioned above, using modular or object-oriented software design is a particularly useful strategy for efficient and robust software verification. In object-oriented design, complex code is built up from simpler atomic components, each of which can function (and, therefore, can be tested) in isolation. Each component has a clear interface that defines its inputs, the modifiable quantities (state), and its outputs, but hides its internal implementations from users (including other components) to help promote compartmentalization and facilitate unit testing. Complexity is engineered through hierarchy, with high-level components coordinating the activities of simpler sub-components. Consider the following example, adapted from Reddy [17], of a multibody system consisting of a collection ("container") of rigid bodies. Suppose the container has two methods: it can return a body's color given its index (e.g., calling the function getColor on "body #2" returns "blue") and it can return a body's index given its name (e.g., calling getIndex on "femur" returns "2"). This design obviates the need to provide a third method that directly returns the body's color given its name, since the two existing methods can be cascaded to do the same job (i.e., calling getIndex on "femur" and then calling getColor on the returned index , "2", will give the desired answer, "blue"). Providing only the first two methods is preferable: if naming or indexing conventions change in the future, the third method will create errors that will go unnoticed unless there is a separate verification test for this new method. Branching logic and circular relationships are also discouraged, even if they make the code shorter, since linear processing is easier to understand, test, and

debug. Components should be reusable so that new and custom code can be minimized, isolated, and easily targeted by tests. The more a block of code is used, the more likely it is to be verified in a wide range of use cases. Much has been written on the topic of object-oriented design (e.g., [17,18]).

More details about verification testing for each part of the modeling and simulation pipeline are included in Section 3. Following the verification stage, you should be confident that your equations (mathematical models) are implemented correctly. You can then begin generating and analyzing your initial modeling and simulation results and proceed with formal validation, described in Sections 2.4 and 2.5, below.

2.4 Validate your results by comparing your model and simulation to independent experiments and other models

The first step of formal validation is comparing the outputs of your model and simulation to as many independent datasets as possible. Validation against independent data is distinct from model calibration (the process of finding the best-fit model and simulation to available data), though both are key steps in any modeling and simulation study. Calibration should be conducted first, followed by comparison to any independent data not used to tune your model or generate your simulation.

A wide range of potential data is available for calibration and validation. These sources include data from a typical motion capture analysis like optical marker trajectories, ground reaction forces, and electromyography (EMG) signals, along with data collected from

imaging or cadaver studies to help define musculoskeletal dynamics and geometry. Additional experimental modalities, such as ultrasound to measure muscle fascicle and tendon dynamics in vivo or instrumented knee replacements to measure internal joint loads, also provide valuable data for calibration and validation. In one common workflow, for example, experimental kinematics and ground reaction force data are used to generate a muscle-driven simulation of the observed motion. In this case, the experimental kinematics and ground reaction force data are used to calibrate the simulation. EMG data are reserved for independent validation to help researchers determine whether the muscle coordination predicted by the simulation is a good fit to experimentally observed timing of muscle activity. Comparing muscle–tendon dynamics to ultrasound data, or predicted joint loads to data from instrumented joint replacements, are additional potential data to aid validation in this example workflow.

We also recommend comparing the predictions of models and simulations to previously validated and published studies, when possible. For example, if you are creating a new simulation framework for studying pathological gait, your framework (minus any pathology-specific changes) should produce gait simulations of normal walking that are a good match to the many walking simulations available in the literature. Data you can compare include joint angles and moments, ground reaction forces, muscle activations and forces (timing and/or magnitude), internal joint loads, and muscle fiber and tendon velocities.

More details about using experimental and previously published modeling/simulation data for calibration and validation of each component of a modeling and simulation framework, including how close a match is expected, are described in Section 3.

2.5 Test the robustness of the study by evaluating the sensitivity of your results to model parameters and other modeling choices

The next step in the validation process is evaluating sensitivity, since validation by comparison to independent experimental data often leaves remaining uncertainty (e.g., one of the model's input parameters may have known experimental measurement error that cannot be eliminated). Many quantities are impossible to adequately measure experimentally (otherwise, we wouldn't need a model!). As a modeler, your key responsibilities are to determine the range of possible outcomes (e.g., confidence intervals) and how sensitive the outputs of interest (i.e., those used to test your hypothesis) are to the input parameters and data—particularly those with large known variability, uncertainty, or influence. Sensitivity analysis can also help determine the level of modeling complexity appropriate for your study by assessing whether your conclusions are robust to a particular modeling simplification or assumption (e.g., reducing the number of musculotendon actuators in your model or ignoring muscle activation dynamics).

Several approaches to sensitivity analysis are available, from simple to complex [19]. The direct approach is a differential analysis, which determines the analytic relationship between inputs and outputs [20]. Although generally not practical for complex, nonlinear

models and simulations, differential analysis serves as the foundation upon which other approaches to sensitivity analysis are built. A straightforward and common approach to sensitivity testing is the parametric study, where the researcher sweeps through a range of input parameters, varying one parameter at a time or multiple parameters in a factorial design, to determine interaction effects. A parametric assessment can be sufficient when possible input ranges and sensitivities are known in advance; however, testing a large number of interactions quickly becomes intractable.

Sampling methods for sensitivity analysis, such as Monte Carlo analysis, are also common in biomechanics research (e.g., [21,22]). The researcher creates random samples of the input parameters given specific probability distributions (e.g., Gaussian with an estimated mean and standard deviation based on data from experiments). The corresponding distribution of the output is then estimated by performing repeated runs of the simulation or analysis. Sensitivity factors are computed based on the relationship between inputs and the output distribution (using a correlation coefficient, for example). The Monte Carlo approach requires many (often thousands) of runs of a simulation and, thus, can become computationally expensive. Various techniques are available to improve sampling efficiency (e.g., Latin hypercube sampling [23]). Modern methods for probabilistic analysis develop an approximate relationship between input and output distributions to substantially reduce the number of simulation iterations required (e.g., using prior knowledge in a Bayesian inference approach [24]). These methods must be validated against the "gold standard" Monte Carlo approach, though initial tests in biomechanics look promising [19]. These approximate methods can provide information about the relative

importance of various input parameters, but do not include enough information to determine the direction of influence.

Several texts and review papers provide overviews of approaches to sensitivity analysis for engineering applications [25,26], including biomechanics [19]. A review of known sensitivities in common modeling and simulation frameworks is provided in Section 3.

2.6 Document and share your model and simulation

After careful validation through comparison to independent data and sensitivity analysis, you will have gained confidence in the ability of your model and simulation to answer the research question you posed; however, the validation process does not end when the last result is computed or figure is generated. An additional, vital step is documenting your modeling and simulation methods, results, and conclusions. In the documentation process, you should clearly indicate how your findings answer your original research question and how your validation process has adequately addressed known sources of error and uncertainty. In most cases, some uncertainty will remain, so you must also describe the known limitations and detail how these limitations might impact your conclusions.

In addition to standard publications in journals, we believe that sharing your models, simulation tools, and results with other researchers and clinicians is essential for validation and helps ensure your research has a broad impact. Allowing others to review your models, simulations, and software can help identify errors and improve your models and simulation tools. As others apply your models and simulation tools to new research

questions and analyses, more information about both the strengths and limitations of your model and simulation will be established. Finally, sharing your simulation data helps expand the pool of available independent data for future researchers to use in the validation process, as shown in the feedback arrow from “Generate simulation” to “Validate your results” in Figure 2.

2.7 Generate predictions and hypotheses that can be tested in the real world

The validation process can continue beyond the life of a single modeling and simulation study by generating hypotheses that you or other researchers test with experimental data. In other words, do the high-level predictions and analyses hold up to independent testing with clinical or experimental data? For example, if a model suggests that plantarflexor muscle strength is vital for maintaining adequate knee extension during stance in children with cerebral palsy, does a plantarflexion strength-training program improve patients’ gait? Does an assistive device to reduce the metabolic cost of uphill walking, designed with a simulation, work when you build the device and test it with human subjects? Establishing more links like these between modeling and clinical or experimental studies is essential for advancing the fields of biomechanics and rehabilitation research.

3 Best Practices for Verification and Validation of Neuro-Musculoskeletal Models and Simulations

In this section, we review best practices for verification and validation of each potential component of a study's modeling and simulation framework (Figure 4). These best practices are based on a review of the literature, standards employed by our research group, and experience working with many researchers using computer models to gain insights into human movement. Since the validation process starts with the formulation of a research question and study design, we begin each section by reviewing modeling choices and providing recommendations, such as adjusting the level of modeling complexity to match that required by the study.

Generation of human and animal movement, and thus computer modeling and simulation of these phenomena, is a complex, multi-step process, as demonstrated by Figure 4. Studies can include models or estimations of neural command, musculotendon dynamics, musculoskeletal geometry, multibody dynamics, and contact and other external forces. The dynamics equations describing these components can be integrated forward in time to generate a simulated movement, or a known motion can be used in an inverse analysis to gain a greater understanding of the muscle coordination and forces involved in generating that motion. A study may include all or only a small subset of the components shown in Figure 4. For example, many studies consider only multibody dynamics, using marker data from a motion capture experiment to determine joint angles. Although this type of motion analysis is often not considered to be modeling, an underlying model—even if the model is

simply a set of bodies, each with six degrees of freedom—is used to formulate equations of motion and calculate accelerations, velocities, and positions over time. Since multibody dynamics is part of nearly all modeling studies, this is where we begin our review.

3.1 Multibody Dynamics

Multibody dynamics enables the description and analysis of interconnected rigid and flexible bodies that move relative to each other, in the presence of forces generated internally and applied externally. Multibody dynamics provides the mathematical formulation (model) of the physical system—the equations of motion—that enable the modeler to calculate accelerations, velocities, and positions over time. A multibody dynamics analysis comes in many forms. We may wish to focus on only motion or kinematics, as in a typical inverse kinematics analysis, or study the motion resulting from forces applied by muscles in a forward dynamic simulation. We can solve for joint reactions to estimate the forces required by ligaments and articulating surfaces in the knee joint to prevent separation and interpenetration. Or, with an induced acceleration analysis, we can determine how much a single force in the system (due to gravity or a muscle, for example) accelerates each body in the multibody system.

3.1.1 Modeling Choices

The first step in performing a multibody dynamics analysis is determining the underlying model that will be used to guide the solution of kinematics and dynamics. The model determines how bodies are permitted to move relative to each other and, in the case of a

dynamics analysis, the distribution of mass in the system (i.e., inertial properties). Every type of motion analysis requires a model, but in many cases, the user of a commercial motion analysis processing package is unaware of the details of the underlying model and whether that model is appropriate for the motion under study. For example, some packages impose no constraints on body motions, which means the model can undergo non-physiological motions like separation or impenetration of body segments [27,28]. *We recommend* performing both inverse kinematics and dynamics analyses using a model that represents physiological joints and is scaled to the anthropometry of the subject. Using a model of the underlying skeletal geometry prevents non-physiological motions and typically makes inverse kinematics and dynamics computations more robust to noise [29]. More details about formulating and validating models of biological joints are included in Section 3.2.

Given an underlying model, the next key decision is how to formulate and solve the corresponding equations of motion. *We recommend*, when possible, using existing multibody dynamics codes that have been designed and verified for engineering and scientific use. Adams (www.mscsoftware.com/product/adams) and Simbody [16] are, respectively, commercial and open source examples. Software that has originally been designed for other purposes, such as gaming and computer animation (e.g., ODE (www.ode.org), Bullet (bulletphysics.org), and Havok (www.havok.com)), must be used cautiously and verified independently, since these codes are designed to optimize speed and appearance rather than physical accuracy.

One common approach to formulating the equations of motion is to use Cartesian coordinates (as in Adams, ODE, and Bullet). Each body segment has six degrees of freedom (i.e., its position and orientation in space), and constraints are added where necessary to model the limitations imposed on the motion by joints. This formulation tends to be easier to implement in software, but requires many nonlinear constraints that are usually solved with approximate iterative methods to maintain tolerable execution times. *We recommend*, instead, using an internal coordinate approach [30], which provides an exact representation of idealized joint motion. The resulting system of equations is small and dense, and is generally solved with exact methods. The internal coordinate approach is more difficult to implement, but several existing dynamics packages such as Simbody [16], MotionGenesis (formerly AutoLev) [31], and SD/FAST [32] contain well-tested implementations.

Once the equations of motion have been formulated, they are solved over time to analyze or predict the motion of a physical system. *We recommend* using error-controlled numerical methods to ensure accuracy and careful state handling, as described by Sherman and colleagues [16]. An important design consideration is the handling of the model's state, which is the set of unknowns or variables that fully describes the system at an instant in time. Tentative state changes are required for error estimation, but can lead to incorrect computations if the state is not handled correctly. Thus, dynamic simulation software must be designed and verified to ensure that tentative state changes persist and desist appropriately.

Additional considerations when choosing a dynamics engine, such as the method for enforcing constraints, the handling of discrete and continuous values, and the user interface for defining accuracy, are included in publications on Simbody [16] and OpenSim [33]. The numerical methods employed by a dynamics engine are also relevant to verification and validation, and are discussed in Section 3.6.

3.1.2 Verifying Multibody Dynamics

Multibody dynamics software can be verified in many ways [34]. *We recommend* ensuring that the physical system adheres to principles of conservation (including mass, momentum, and energy) across a wide range of conditions. For example, a common verification task is ensuring that constraints and ideal joints do no mechanical work. In a forward dynamics simulation, all applied forces can be divided into their constitutive power generators (ideal actuators), storage components (e.g., potential energy due to gravity or elastic strain), and dissipative components (e.g., friction or viscosity). In this case, verification should confirm that system energy minus work plus dissipated energy is conserved for each force and actuator, and also for the multibody system as a whole.

We can also represent the same physical system using different formulations, the comparison of which is another avenue for verification. For example, joints can be modeled with internal coordinates or free joints with constraints. These formulations should produce identical solutions (to within numerical tolerance). When developing new multibody simulation software, *we recommend* verifying that a modeled system in your code produces the same results as the identical system modeled in existing code(s).

3.1.3 Validating and Evaluating Robustness of Multibody Dynamics

Following verification, a multibody dynamics engine can be used to calculate motions and forces for the system of interest. The accuracy of your multibody dynamics solution is limited by the quality of the input data you provide—typically from motion capture and force plate recordings—and the quality of the underlying model. Traditional marker-based motion capture systems generally have system errors of 1–5 mm [35], which is combined with soft tissue artifacts of up to 10 mm for human movement [36]. Joint torques calculated from inverse dynamics have been shown to be sensitive to both force plate recordings and kinematics [37,38]. Thus, motion capture systems must be regularly and carefully calibrated and you should clearly document all steps of your experimental protocol (e.g., with written notes, photos, and videos). Errors in the underlying model can also create discrepancies, particularly when models of the joints are not included. Careful model selection and subject-specific scaling, as discussed in Section 3.2.1, are essential for obtaining multibody dynamics models that closely match closely to experimental measurements. We illustrate practical validation of multibody dynamics by returning to our case study on running. The case study was introduced in Figure 3; in Figure 5, we review the process the investigators used for validating multibody dynamics, the first step of which was careful data collection and model scaling.

We recommend comparing calculated joint angles and moments to those reported by previous studies, when possible. For example, kinematics and kinetics for walking and running at multiple speeds for healthy individuals are readily available in the literature

(e.g., [12,39]) and new simulation results should be within 2 standard deviations of previously reported values. Researchers often use a model to reproduce experimental kinematics from external markers placed on a subject. *We recommend*, in this case, testing that model marker locations fall within the measurement accuracy (combing calibration accuracy [35] with skin movement [36] or marker placement [40] errors) for each marker.

An additional approach to validation is comparing the acceleration of the system center of mass to the net externally applied force (i.e., does $F=ma$?), as demonstrated in the case study of Figure 5. In many motion-capture assessments of inverse dynamics, the size of the discrepancy between motion and forces is neither calculated nor reported—a dangerous omission, since a discrepancy nearly always exists [41]. In a motion capture experiment, kinematic data (obtained from optical markers or inertial measurement units) are measured independently from kinetic data (obtained from force plates and sensors) and errors in either of these data can lead to inconsistencies.

Dynamic inconsistencies should be calculated using the model's equations of motion by comparing the net forces and moments acting on the defined center of mass (COM) from the measured ground reaction force to the product of the system's mass matrix and COM accelerations calculated from inverse kinematics. When forward simulations are performed (e.g., in Computed Muscle Control [42]) or a force balance is calculated (e.g., in Static Optimization [43]), residual or "hand of god" forces are typically applied to achieve dynamic consistency between the motion and the applied forces. In general, *we recommend* ensuring that residual forces and moments are small enough that they do not affect the

conclusions of the study (e.g., if a residual moment is contributing significantly to the accelerations of a variable of interest, the residual must be reduced). In particular, *we recommend* force discrepancies that are 5% or less (peak and RMS) than the magnitude of the experimentally measured net external force and residual moments that are less than 1% of COM height times the magnitude of the measured net external force.

The first key step in minimizing dynamic inconsistencies is well-calibrated data collection and careful processing, along with proper scaling of the model to the subject of interest, as described above. If $F=ma$ discrepancies are still large, additional adjustments can be made to the underlying model (e.g., changing inertial parameters) or the kinematics (e.g., eliminating motions that the model cannot achieve) to reduce or eliminate the dynamic inconsistencies [6]. If the recommended agreement cannot be achieved with reasonable adjustments to the model or kinematics, the assumed model may be of insufficient quality to capture the underlying system dynamics or to represent the observed experimental data. For an example of this process, see the case study of Figure 5, where the investigators ensured that residual forces and moments were reduced to the recommended thresholds, while maintaining kinematics that were within measurement error..

3.2 Musculoskeletal Geometry

The next layer of detail that many studies include is a model of muscle and skeletal geometry. Musculoskeletal geometry includes properties of bodies, which for the purposes of this paper are rigid with specified mass and inertial properties. The definitions of joints, the connections between bodies, are also part of musculoskeletal geometry. Joints can

range from generic mechanical connections, like pins and gimbals, to more specialized relationships that define the coupled motions and complex constraints common in biological joints. Passive structures like ligaments and cartilage can be modeled explicitly or with lumped passive joint properties. A final component is the geometry of muscle–tendon units, including their attachments to bones and wrapping over and through other structures, such as deep muscles and retinacula. Bone, muscle, and joint geometry together determine muscle moment arms, which map muscle forces to joint moments. *We recommend* choosing anatomically based models of musculoskeletal geometry that represent physiological joint kinematics and muscle path geometry if you are interested in studying internal joint loads or the accelerations produced by muscles during movement.

3.2.1 Modeling Choices

Modelers must balance the need to capture the features of the biological system and motion of interest with the added computational cost and validation requirements of increasing model complexity. *We recommend*, in general that you simplify musculoskeletal geometry where possible and focus on accurately representating the muscles and joints under study. For example, when studying a primarily lower extremity motion like walking, you can typically simplify the upper extremity by lumping body segments and muscles or actuating the upper extremity with joint torques rather than muscles. Conversely, for primarily upper extremity motions, the lower extremity can often be simplified, allowing more detail to be included for the muscles and joints in the trunk, arms, and shoulders. Another common choice is whether to use a planar or fully three-dimensional model. Cycling, walking, and running are principally sagittal plane activities, so two-dimensional

models can be appropriate for studying gross motor coordination and energetics in these cases. Nevertheless, out-of-plane motions and muscle activity often play a role in research questions related to pathological gait or the stability of locomotion, and should be modeled in these situations. These modeling choices are challenging and in cases where the appropriate level of modeling complexity is unclear, *we recommend* careful validation against independent data and sensitivity testing (e.g., to determine how reducing the number of muscles in your model or simplifying joint geometry affects the outputs of your simulation).

Bodies

The dimensions and inertial parameters of the rigid bodies that compose a human or animal model have been measured in cadavers and via imaging in living subjects to create a generic set of anthropometry (e.g., [44,45]) that is then scaled to represent individual subjects. The anthropometry of bodies can also be estimated on a subject-specific basis with dual-energy X-ray absorptiometry (DXA) scans, magnetic resonance imaging (MRI), and 3D body scans (e.g., [46–48]). The dimensions of a subject's body segments can also be estimated from manual measurement or from markers by finding the distance between functional joint centers (e.g., [49,50]) or between markers placed on anatomical landmarks.

We recommend, at minimum, obtaining marker-based or careful manual measurements of segment lengths and using these measurements to scale the dimensions and inertial properties of a generic model to an individual subject. If feasible, imaging techniques can provide valuable additional information in the scaling process. *We recommend* subject-

specific modeling through MRI, measurements of bone alignment, or other means in cases where body segment geometry is known to deviate substantially from existing generic models [51].

Joints

In Section 3.1, we discussed how to map the description of a joint's motion to a system of dynamic equations. In this section, we review how to define a joint's motion based on experimental measurements. In the musculoskeletal system, the motion permitted by a particular joint depends on the geometry and properties of the cartilage and bone at the articular surface, along with other structures like menisci, the joint capsule, ligaments, and muscle (e.g., the rotator cuff muscles of the shoulder). The effects of these various structures can be represented in a joint model in several ways, from simple to complex, including less or more anatomical detail.

On the simple end of the joint modeling spectrum, skeletal joints can be built from basic mechanical joints (e.g., a ball-and-socket joint is a reasonable approximation of the hip), often with passive elements or other constraints to limit the range of motion and represent the net action of passive structures (e.g., to prevent excessive hip extension beyond physiological ranges). Constraints can also be used to represent coupled motion. For example, one common model of the knee joint couples the translation of the tibia with respect to the femur as a function of knee flexion angle [52,53]. At the complex end of the spectrum, a joint can include explicit contact and ligament models that constrain the joint to follow physiological motion. This type of model allows more detailed analysis of loads,

but incorporates many parameters that are potentially difficult to measure and must be calibrated and validated. For many applications, an explicit contact and ligament model at the joint is unnecessary. Models combining simple mechanical or custom biologically-inspired joints with constraints and lumped passive properties are typically sufficient to reproduce gross motions like walking, reaching, running, cycling, etc. These model formulations can also provide net internal joint moments (e.g., [53]).

We recommend, as discussed above, using internal coordinate formulations that permit motion in only the known degrees of freedom, rather than models with many degrees of freedom and many constraints. Avoiding the use of constraints is particularly important for models that include body segments whose masses are small relative to the applied forces (as in the upper extremity, for example). Incorporating constraints in such models is computationally expensive due to the numerical singularities approached as segment masses approach zero.

Several types of experimental data are available to help formulate joint models. In cadaver experiments, the joint of interest can be isolated and its kinematic and dynamic properties can be measured externally with a loading rig (e.g., [54–56]. High-resolution imaging and bone pin data can also provide information about the relative motion of bodies in the skeletal system [57,58], as can less invasive—though less accurate—surface marker motion. *We recommend* basing new joint models on high-quality data from cadavers, bone-pins or high-resolution imaging, including cinefluoroscopy [59] which provides high-speed, high-resolution data of loaded joints.

When using existing joint models, *we recommend* that you acquire motion capture data to help determine subject-specific joint geometry (often called functional joint centers). Identifying joint axes based on anatomical landmarks can be problematic, since there is significant variability in marker placement between and within examiners [40]. Ehrig and colleagues [49] and MacWilliams [50] review methods for finding functional joint centers. In cases of pathology, MRI data can also provide detailed subject-specific information about joint center locations [51]. When choosing a joint model, you must also ensure the model has been validated for the entire range of motion you plan to study.

Muscle Geometry

The geometry of muscle–tendon units can include detailed path information, or the modeler can simplify muscle geometry with regression or other equations that directly define moment arms as functions of joint angles. For muscles with broad attachment sites (e.g., the gluteus medius), several muscle segments may be required to capture the varying moment arms of the muscle. Whether using moment arms or path geometry, a key modeling choice is the number of muscles used in the model; you should tune the model's complexity to suit your research question. Studying the contribution of the abdominal muscles to stability during walking, for example, requires more detailed trunk musculature than studying the contributions of the ankle plantarflexor muscles.

Muscle moment arms can be measured in cadavers directly using tendon excursion experiments or using load experiments (e.g., [60–62]). Muscle moment arms can also be

estimated experimentally using MRI, computerized tomography (CT) scans, or the digitization of cadavers to determine a muscle's line of action and the perpendicular distance to the joint center of rotation (e.g., [61–63]). *We recommend* using moment arm data to calibrate and validate your musculoskeletal model, as discussed in Section 3.2.3.

Detailed models of muscle geometry include muscle paths defined by origin and insertion points, which are often combined with via points and wrapping surfaces to represent constraints from retinacula and to prevent muscles from passing through bones or reaching other non-physiological configurations. *We recommend* modeling detailed muscle path geometry to accurately predict muscle–tendon dynamics and internal joint loads, although this approach is more computationally expensive than using pre-computed moment arms. Muscle geometry can be determined by dissecting and digitizing muscles in cadavers (e.g., [64–66]) or can be estimated by segmenting MR images (e.g., [63,67]). Origin and insertion points are typically defined as the centroids of the muscle attachments in the cadaver or imaged subject, but via points and wrapping surfaces have been defined several ways. For example, Horsmann and colleagues fit geometric shapes to measured points on the surface of a cadaver [65]. *We recommend* using wrapping surfaces and via points to achieve physiological wrapping over bones and to match experimentally measured moment arms (e.g., [7,68]).

Modeling Assumptions and Limitations

Many phenomena are often ignored or simplified in models of musculoskeletal geometry. The geometry of a muscle is typically reduced to one or more line segments. This

simplification makes simulation computationally tractable, but fails to represent the true underlying complexity of muscle geometry. Further, multiple muscles can attach to a single tendon (e.g., the triceps surae), but models typically assume separate tendons exist for each muscle component. The deformability of bodies is also ignored. Some parts of the human musculoskeletal system are nearly always drastically simplified. For example, there is no widely used and well-validated detailed model of the foot and connective tissues (e.g., the iliotibial band) are often ignored or simplified. When documenting results, these limitations should be recognized and researchers should discuss potential impacts on the study conclusions. These assumptions are also important areas for future research to improve current models.

3.2.2 Verifying Musculoskeletal Geometry

The key steps for verification of skeletal models are described in Section 3.1.2. *We recommend*, in short, verifying that bodies and joints adhere to the laws of physics. Given the complexity of human and animal musculoskeletal geometry (e.g., joints may be complex and muscles may include via points and wrapping surfaces), *we recommend* using the effective torque method to calculate muscle moment arms, as described by Sherman and colleagues [69]. To verify muscle geometry code using the effective torque method, the model developer applies a unit tension to the muscle of interest and uses the system's equations of motion to calculate the resulting (effective) moment about the joint; the moment arm is the moment divided by the unit muscle force. The moment arm computed using the effective torque method should be verified against the moment arm (r) calculated as the change in muscle-tendon length ($d\ell$) resulting from a change in the joint angle ($d\theta$),

or, $r = d\ell/d\theta$. As observed by Sherman et al., [69] this equation is a consequence of the assumption that all related constraints are workless and is commonly used to measure muscle moment arms experimentally (known as the "tendon-joint excursion method) [62].

3.2.3 Validating and Evaluating Robustness of Musculoskeletal Geometry

Each component of musculoskeletal geometry should be validated, where possible, along with the musculoskeletal model as a whole.

Bodies

Anthropometry varies by age, sex, pathology, and natural variation (e.g., [44,45,70,71]). Errors in model dimensions and inertial parameters can affect joint angle and moment calculations [72–74], which will result in errors in estimated muscle forces. If performing a dynamic analysis, *we recommend* comparing the whole-body center of mass acceleration computed from kinematic data to the acceleration generated by measured ground reaction forces. If there is a large discrepancy, one possible source of error is missing or poorly modeled body segments (see Section 3.1.3 for more details).

Joints

Joint properties (e.g., centers of rotation and laxity) vary between individuals, particularly in the case of injury or pathology (e.g., [75]), and can have a significant effect on joint angles, moments, and internal loads (e.g., [76]), as well as muscle moment arms (e.g., [51]). In a Monte Carlo–based study of simulated maximum thumb tip forces, Valero-Cuevas and

colleagues [22] identified the kinematic description of the joint as being the most influential parameter preventing the model from generating physiological maximum forces. When developing a joint model, *we recommend* ensuring that modeled kinematics are within measurement error of high-resolution experimental data (e.g., from cinefluoroscopy or bone pin data). When using the model in a simulation, the dynamic inconsistencies between the model, kinematic data, and force data should also be small as described in Section 3.1.3. Poor joint modeling is a common source of error. Further, *we recommend* comparing predictions of internal joint loads to estimates from instrumented joint replacements; several datasets are available for the hip and knee joints [77–80].

Muscle Geometry

Muscle geometry and moment arms vary between individuals as well. For example, Murray et al. [66] found that moment arms could vary by a factor of two, although these differences could be accounted for by using a simple scaling technique. Herzog and Read [81] also observed considerable difference in moment arm magnitudes but found that the variation of moment arms with joint angles were generally similar between subjects. Duda and colleagues [82] digitized femoral muscle attachment sites for six cadavers and found that the standard deviation of centroid locations was 80% of the mean. Attachments can also vary with pathology, such as bone deformities common in children with cerebral palsy [67].

Given this variability and uncertainty, careful validation is an important step in any study that includes muscle geometry. When using a model of muscle geometry, *we recommend*

comparing the model's moment arms to experimental data to ensure that magnitudes and trends are similar (ideally within 2 standard deviations) throughout the range of motion that you plan to study. You should also compare modeled muscle paths to imaging data (e.g., from MRI). Further, *we recommend* that you combine musculoskeletal geometry with muscle-tendon dynamics and perform indirect validation by comparing simulated joint moments to experimentally measured values, again throughout the range of motion of interest. Validation data can include maximum moments from dynamometer testing, experimentally measured passive moments through a range of motion, and inverse dynamics–based moments from motions like walking, running, or reaching. In our research group, we aim to achieve moments within 2 standard deviations of experimental values, unless known experimental uncertainty or documented modeling assumptions can account for additional discrepancies. In cases where agreement is not achieved, you must clearly document the limitations of your model and take these limitations into account when interpreting the results of your simulation. In the case study of Figure 6, we demonstrate the process of comparing the outputs of a musculoskeletal model to experimental moment arms and passive and active net joint moments.

Muscle–tendon kinematics, dynamics, and forces (and, thus, internal joint loads) are all sensitive to a model's muscle geometry and moment arms. *We recommend* performing a sensitivity analysis when your research question depends on muscle geometry or moment arms, such as when the question demands estimating joint loads or muscle force magnitudes. Several investigators have explored the sensitivity of predicted muscle forces and activations to moment arms and muscle geometry. In general, the magnitude of a

muscle force is more sensitive to geometry, while timing of muscle activity is fairly robust to changing muscle geometry [83–86]. Correa and colleagues [86] studied muscle contributions to COM and joint accelerations using models with different moment arms, and found the predicted muscle function to be similar, though force magnitude varied. Grouping muscles with similar paths was found to have little effect on the timing or coordination of estimated activity in a muscle-driven simulation of walking [87]. There is evidence that joint force predictions are sensitive to how muscles are discretized into paths in a given model [88], but more work is needed to understand this relationship.

3.3 Muscle–Tendon Dynamics

While the muscle geometry described above can characterize the motion of the whole muscle–tendon unit, a model of muscle–tendon dynamics is required to estimate muscle activations, muscle forces, and the interplay between muscle and tendon during motion. In this section, we focus on the most widely used computational model of muscle, the Hill-type model, which includes a contractile fiber, series and parallel elastic elements, and activation dynamics.

3.3.1 Modeling Choices

Modeling muscle–tendon dynamics adds significant complexity as well as many additional parameters to calibrate and outputs to validate. In some cases, it is reasonable to ignore or simplify muscle–tendon dynamics. Examples where muscle–tendon dynamics can be ignored include studies where only net joint moments or powers are required, or in purely

kinematic studies of joints or whole muscle–tendon units. Inverse dynamics with static optimization to resolve muscle forces frequently ignores tendon compliance and enables estimation of muscle forces and joint loads. *We recommend* including a full computational model of muscle and tendon dynamics when your research question depends on the interaction between muscle and tendon during a motion or when elastic storage of energy in tendon is a known or likely contributor to the motion of interest (e.g., high-force motions like running). Further, recent work in predictive simulation has shown that muscle–tendon dynamics are required to achieve human-like motion [89,90]. If computational or other restrictions require making simplifications, *we recommend* using sensitivity testing to determine the impact of omitting complexity (e.g., modeling rigid vs. compliant tendons) in your model.

Muscle–Tendon Dynamics Input Parameters

Before validation, you must understand and calibrate the parameters of the muscle–tendon dynamics model. Parameters for Hill-type muscle models include the maximum isometric force of the muscle, determined by combining estimates of physiological cross sectional area (PCSA) with muscle specific tension; pennation angle; the force–length–velocity relationship; the passive muscle force-length relationship; timing parameters for activation dynamics; and the stiffness and slack length of the tendon. Architectural parameters (e.g., pennation angle) are typically derived from measurements in human cadavers, while the parameters that define dynamic muscle force generation (e.g., the force-length-velocity relationship) are typically derived from isolated muscle experiments in animals. Additional data can also come from MR images (to determine muscle volumes), laser diffraction (to

determine sarcomere lengths), and ultrasound imaging (to determine muscle fascicle lengths).

We recommend creating models with muscle-tendon dynamics based on comprehensive muscle and tendon architecture data from cadaver and MRI measurements of multiple subjects. Models using MRI-based estimates of muscle PCSAs (e.g., [91]) are preferred, since cadaver measurements are typically obtained from older subjects whose PCSAs can be substantially smaller than those of young, healthy subjects [92–94]. Many current models are built from the older Wickiewicz dataset [95], which was obtained from a small number of cadavers without the aid of laser diffraction to measure sarcomere lengths. The fiber lengths in this study differ by 10–100% from the more recent Ward dataset [96]. The optimal fiber lengths in the dataset provided by Ward and colleagues are based on more subjects as well as measurements of sarcomere lengths.

Hill-type muscle models rely on normalized curves for the force–length, force–velocity, and passive properties of muscle fibers based on experiments in rat, cat, and rabbit muscle, where force output is measured for a range of lengths and activations levels (see [97] for a review). *We recommend* using muscle model curves that are a close match to available experimental data (e.g., [98]).

Simplifying Muscle–Tendon Dynamics

It may be reasonable to simplify muscle–tendon dynamics, depending on the research question and outputs of interest. Some studies ignore the series elastic element and assume

the tendon is inextensible, which can significantly decrease computation time [98]. Assuming a rigid tendon is valid for analyzing muscles with short or very stiff tendons (e.g., some of the upper extremity muscles and muscles crossing the hip) or in slower motions with low forces in which tendon stretch is minimal (e.g., walking [99]). The parameter that affects the tendon modeling decision is the ratio of the muscle optimal fiber length to the tendon slack length [64]. *We recommend* including tendon compliance in motions with large forces like running or when analyzing elastic energy storage in tendons, where compliance plays a crucial role. For example, in an optimal control simulation of running, ignoring tendon compliance reduced the maximum sprinting speed achieved [100].

Additional simplifications include ignoring the force–length–velocity relationship or activation dynamics, although *we recommend* including these properties, in most cases. Excluding the force–length–velocity relationship did not significantly change the muscle forces during walking predicted by static optimization [99], although many muscles have been shown to have a wide operating range on their force-length and force-velocity curves even during walking [9]. Ignoring activation dynamics also had a small effect on static optimization solutions of walking [99], although its exclusion reduced maximum simulated running speed in a dynamic optimization study by Miller and colleagues [100]. Activation dynamics must also be included when using experimental EMG measurements as the neural control input [8,9].

Assumptions and Limitations

Several phenomena are ignored in Hill-type muscle models. Experimental evidence exists for variable-gearing pennate muscles [101], force enhancement [102], short-range muscle stiffness [103], and variations in the force–length curve at submaximal activation [104], but these phenomena are not accounted for in typical Hill-type models. The mass of muscles and effects of fatigue are also typically ignored, as is the distribution of fast- and slow-twitch fibers within a muscle. Fiber lengths and velocities are assumed to be constant throughout the muscle, in contrast with recent imaging studies and finite element models [105,106]. In addition, properties of tendon (e.g., plasticity and creep) are typically omitted from musculotendon models. These limitations should be acknowledged when documenting your findings, and represent valuable areas for future work, combining experiments with models and simulations.

3.3.2 Verifying Muscle–Tendon Dynamics

Before performing validation, *we recommend* the following set of verification tests for muscle models. You should verify that the active and passive fiber force–length, fiber force–velocity, and tendon force–length curves are C2 continuous and that all force-length curves have no negative stiffness regions (other than the descending limb of the fiber force–length curve). These criteria prevent failure of numerical methods during simulation, such as encountering numerical singularities, and prevent nonphysical phenomena (e.g., collagen under strain does not exhibit negative stiffness). You should also check for self-consistency between the action of the whole muscle and that of the constitutive passive and active

muscle components (Figure 7). Further, energy should be conserved in the muscle–tendon unit (i.e., fiber work minus fiber strain potential energy should always equal tendon work minus tendon strain potential energy) during a wide range of conditions. To verify activation dynamics, you can input a step excitation and ensure the activation signal has the characteristic first-order rise or fall time (assuming a first-order activation dynamics model) and remains between the muscle model's minimum and maximum activation values.

3.3.3 Validating and Evaluating the Robustness of Muscle–Tendon Dynamics

Validation of muscle–tendon dynamics can be accomplished through several means. First, when creating a new computational muscle model, *we recommend* comparing muscle forces predicted by a simulation to independent, experimentally measured forces and dynamics (i.e., length and velocity profiles) from isolated muscle experiments for a range of activation conditions (e.g., [98,107]). Millard and colleagues [98] reported mean absolute force errors of 9% for maximal activation tests and 16% for submaximal activation tests. A similar or better match for new models is desirable.

When using an existing muscle model to study a new subject or a new type of motion, *we recommend* combining the muscle–tendon dynamics model with a model of musculoskeletal geometry to transform calculated muscle forces into joint moments, then comparing simulated moments to experimental and inverse dynamics moments as described in Section 3.2.3 and illustrated in the case study of Figure 6. Ultrasound data can

also provide information to help validate muscle fascicle and tendon lengths and velocities for muscle–tendon units during motion (e.g., [108]).

Muscle–tendon dynamics predictions depend on the error, uncertainty, variability, and sensitivity of each parameter that defines the computational muscle–tendon model, as described in the following sections. When determining sensitivity of muscle–tendon models, defining the output of interest is essential. For example, single-muscle experiments are generally more sensitive to muscle model parameters than whole-body walking simulations, where forces are generated to track a specified trajectory or meet walking goals [109]. Timing of muscle activity is also generally more robust than predicted force magnitude, at least for walking and running simulations [83,110]. Little is known about the sensitivity of joint reaction forces to muscle–tendon model parameters, though joint forces are known to be sensitive to the distribution of muscle forces (e.g., [111]).

Maximum Isometric Force

The maximum isometric force of a muscle, derived from measurements of specific tension and PCSA, is a key parameter of computational muscle–tendon models. Specific tension has been estimated in several experiments for humans and animals [112,113] by determining the ratio between maximum measured muscle tension and a known PCSA, but a large range of values (11-47 N/cm^2) has been reported.

PCSA can be estimated by measuring muscle mass in cadavers or measuring muscle volumes with MRI. In adults, PCSA typically scales uniformly with the product of height and

mass in the lower extremity [91] and with overall muscle volume in the upper extremity [114]. In many past studies where muscle PCSAs were based on cadaver data, the specific tension parameter was selected to inflate muscle strengths. Using MRI-based PCSAs with known size relationships should avoid the need to artificially increase the specific tension parameter to inflate muscle strengths. The relative distribution of muscle strengths can change as a function of age in the upper extremity [115] and lower extremity [116], though linear scaling laws hold among individuals of similar age. MRI-derived volumes will allow development of improved age-matched musculoskeletal models. More work is needed to understand how muscle volumes and strengths are affected by pathology.

The predicted patterns of muscle force and activation for many simulations are not especially sensitive to the maximum isometric force [109,110,117], particularly when timing rather than magnitude is the most important simulation output [84] and the relative strength distribution between muscles matches experimental measurements. *We recommend* avoiding non-uniform muscle strength scaling, unless isometric force values are based on subject-specific experimental data.

Pennation Angle

Pennation angle is estimated by measuring muscle fiber orientation with a goniometer in dissected cadaver specimens or in vivo with ultrasound. The pennation angle of a muscle can vary within a muscle and throughout a motion. For example, the gastrocnemius pennation angle varies 20–30 degrees during running [108]. Within the muscle, the standard deviation of the pennation angle in gluteal and soleus muscles is large, nearly

equal to the measured value [96]. Fortunately, muscle force solutions seem to be less sensitive to pennation angle than other muscle model parameters [117]. Further, in ultrasound experiments, tendon strain computations were not strongly sensitive to pennation angle [108]. More work is needed to understand how pennation angle affects predictions of muscle fiber and tendon lengths, velocities, and strains during motion.

Optimal Fiber Length

The optimal fiber length parameter is estimated by measuring fiber lengths in dissected cadaver muscles and, in more recent studies, by combining these fiber length measurements with laser diffraction–based sarcomere length estimates. Optimal fiber lengths must be carefully examined in the calibration and validation process. Optimal fiber length varies by subject and within regions of each muscle [96] and muscle force predictions have been shown to be sensitive to the optimal fiber length during gait [109,117–119], and likely most other motions.

F–L–V properties and passive stiffness of muscles

The animal-based data that help define normative force–length and force–velocity curves are assumed to scale well to human muscles, though direct validation of this assumption is difficult. Imaging and finite element models indicate that not all fibers within a muscle operate at the same length or velocity for a motion [105,106], which could have the effect of making whole-muscle force–length and force–velocity curves broader than the single-fiber curves.

Activation Dynamics

Activation dynamics are typically modeled as a first-order differential equation that relates the rate of change of muscle activation to the current muscle excitation. Single-muscle and walking simulations are not particularly sensitive to changes in the activation/deactivation time constants [109].

Tendon Properties

The key properties of tendon that must be measured for use in simulation are its stiffness and slack length. Butler and colleagues [120] reported Young's Moduli of 362 MPa and 613 MPa for human cadaver gracilis and semitendinosus tendons; however, the physical properties of cadaver tissue can be affected by storage method, donor age, and other factors. Maganaris and Paul [121] measured human gastrocnemius tendon stiffness in vivo using real-time ultrasound imaging, and reported a Young's Modulus of 1.16±0.15 GPa. These results agree with analogous in vivo measurements of the (less highly stressed) human tibialis anterior tendon [122], suggesting that tendon material properties are independent of function.

Tendon slack length cannot be easily measured in cadaver or imaging experiments. Researchers have typically measured fiber lengths and joint positions in cadaver experiments. Then, when building a model, slack lengths are set such that the fiber lengths in the model match the measured fiber lengths at the given position (e.g., [7]). It has also

been shown that, of all the musculotendon model parameters, tendon slack length has the largest effect on predictions of muscle forces [85,109,118,119]. Changing the tendon slack length alters where muscles operate on the force–length curve, affects the joint angle where peak force is generated, and changes the range over which a muscle can generate force. Muscles with a high ratio of tendon slack length to optimal fiber length (e.g., greater than three) are most sensitive (e.g., gastrocnemius and rectus femoris) [64]. *We recommend* testing the sensitivity of your research question to tendon properties, particularly when your results depend on muscles whose tendon slack lengths is large relative to the muscle fiber length. An example sensitivity analysis for tendon compliance is shown in the case study of Figure 8.

3.3.4 Energetic or Metabolic Modeling

Muscle activation is the primary driver of metabolic energy consumption during motion and is often a good surrogate for energetic cost. However, computational models of muscle energetics have the potential to provide more information about efficiency and energy use for the whole body and individual muscles during a range of motions. Several energetic models are available [123–126], which use muscle activation along with muscle length, muscle velocity, excitation and/or fast- and slow-twitch fiber composition to estimate the combined mechanical work done and heat released by muscles. These models are based on experiments with isolated fiber bundles from mouse and frog muscles. These models predict whole-body metabolic cost for walking and running with reasonable accuracy, but predictions are sensitive to the metabolic model, corresponding muscle model, and neural controller [90,127]. This is an important area for future work as interest in metabolic

energy and efficiency grows in the community (e.g., in pathology or for designing assistive devices).

3.4 Contact and Other External Forces

In many simulations, experimentally measured contact forces and torques are applied to the model. These measurements can occur either before the simulation (e.g., ground reaction forces collected in a laboratory [11]) or during the simulation (e.g., in haptic feedback applications [128], where the simulation trajectory is not known a priori). In both cases, the quality of the contact force signal is limited only by the quality of the measurements; however, both strategies require experimental equipment and expertise. Contact models are useful when experimental data would be difficult or impossible to collect, such as when modeling internal joint loading [129].

3.4.1 Modeling Choices

We recommend choosing the contact model of least complexity and greatest computational efficiency that captures the phenomena of interest. In its simplest and most efficient form, contact can be idealized as a set of one or more kinematic constraints; alternatively, the modeler can employ a rigid or elastic contact model. Kinematic constraints represent interactions between infinitely stiff bodies, and are satisfied to within numerical tolerance by arbitrarily large contact forces. Ideal constraints such as those representing pin and slider joints do not exist in reality; however, computationally efficient algorithms exist for simulating the resulting differential-algebraic equations. Thus, the constraint-based

approach is useful when you don't need information about detailed interactions, as is the case when studying muscle coordination in cycling, for example [130]. Constraint-based contact models can also reproduce ground reaction forces in walking and running when kinematics are known. In the case study of Figure 9, we illustrate the process of validating a constraint-based contact model when performing an induced acceleration analysis.

An elastic foundation model represents an array of independent linear or nonlinear springs that generate force as a function of penetration depth of one surface into another. Such compliant contact models are convenient and can produce accurate results when properly tuned, but they can be computationally expensive when the interacting materials are stiff. Rigid contact models address this limitation by approximating the behavior of compliant models, treating impacts as instantaneous events and simulating persistent interaction using constraints. Although rigid contact models can be more efficient than their compliant counterparts, they are less realistic—in some cases, even non-physical [131].

Assumptions and Limitations

As modeling and simulation of humans in interaction with external devices (e.g., prostheses or exoskeletons) become more common, additional models and experimental data are needed to understand and validate the complex contact that occurs at the device-soft tissue interface. More work is also needed to develop foot-ground contact models that are suitably fast for predictive simulation while still predicting contact forces that are in good agreement with experiments.

3.4.2 Verifying Contact Models

To verify the simulation of contact, *we recommend* checking that impacts are always energetically neutral or dissipative, normal forces are always repulsive, sliding friction is always dissipative, and other physical laws (such as those of Newton and Coulomb) are respected.

3.4.3 Validating Contact Models

Validation tasks are guided by the research question and application. If studying the upper extremity, for example, the simulation of foot–ground contact is likely irrelevant. In other cases where ground contact plays a significant role in the motion of interest (e.g., studying the triceps surae using predictive simulations of running), *we recommend* validating the foot–ground contact model against experimentally measured ground reaction forces to be certain that appropriate stiffness and damping parameters or an appropriate constraint type (e.g., see the case study of Figure 9) have been selected. Similarly, data collected from instrumented knee joint replacements can be used to tune or validate contact parameters for studying internal joint forces. Finally, if using a rigid contact model, *we recommend* validating its performance against the behavior of the compliant model it has been designed to approximate.

3.5 Neural Control

Neural control coordinates the muscle forces that drive motion and is arguably the most challenging component in Figure 4 to model. Neural commands originate in the brain to

execute planned motions, and combine with feedback from the peripheral nervous system (e.g., stretch feedback from muscle spindles). Thus, neural control of movement arises from a complex interaction of neural commands descending from the brain and reflex responses and sensory information delivered to the central nervous system from a vast set of sensors located in muscles, tendons, joints, and skin.

3.5.1 Modeling Choices

In some cases, a model of neural control is not required, since many biomechanics studies depend only on joint angles and net moments. Other studies use experimental input from EMG to estimate the neural command and drive motion or muscle dynamics. EMG-driven simulations do not generally produce a forward simulation that generates a coordinated movement unless additional, corrective forces are applied to maintain the desired trajectory. It is also difficult to acquire EMG for some muscles in the body. Nevertheless, the EMG-driven approach has been successfully used to understand muscle–tendon dynamics in walking and running (e.g., [8,9]) and to simulate various upper and lower extremity motions (e.g., [132,133]).

Modeling of neural control generally comes in two forms. The first, tracking control, is applicable when the dynamics of a motion are known (e.g., from a motion capture experiment) and the controller is responsible only for solving the muscle redundancy problem to resolve a net joint moment into a constituent moment from each muscle. The second form of neural control modeling, predictive simulation, is applicable when the motion is neither known nor assumed and the controller must represent the goals of a task

to be performed (e.g., maximize jump height or minimize cost of transport) and optimize control signals to find the motion trajectory along with the moments or forces required to obtain that motion.

For both types of neural control modeling, an effort-based cost function is often used. When simulating walking and running, the specific muscle-based cost function chosen (e.g., energy expended, activation squared, and activation cubed) does not have a large effect on the overall muscle coordination strategy [90,134], though it affects predicted internal joint loads [111] and estimated cost of transport [90]. While the commonly used effort-based objectives show good results in many applications, their suitability in pathological or maximum performance motions (e.g., sprinting or jumping) may be less appropriate. For example, in individuals with osteoarthritis, muscle coordination might be tuned to minimize a combination of effort and joint loads. Tracking simulations have also been shown to over-predict the cost of transport due to co-contraction when precisely matching measured motions [90]. In these situations, validation against experimentally measured muscle activity and internal joint load measurements (from instrumented joint replacements) is important.

The majority of current studies that include a neural control model are concerned only with solving the muscle redundancy problem to track experimentally measured motion. An effort-based objective (e.g., minimizing the sum of squared muscle activations) is used to solve for a set of muscle excitations or forces that track measured motions within a specified tolerance. Muscle forces can be resolved at each time step during the motion of

interest (i.e., using static optimization) or over the entire motion trajectory (i.e., using dynamic optimization). The effort minimization approach, whether in a static or dynamic optimization, has been used to understand muscle coordination in a range of motions, including walking, running, cycling, and reaching (see [135] for a review). In walking, static and dynamic optimization yield similar predictions of muscle and joint forces [99].

There is also growing interest in using predictive simulation to synthesize motions *de novo*. In predictive simulation approaches, a high-level motion task is specified and control signals or controller parameters are determined to achieve this task. This approach has been successfully used to simulate maximum-performance activities like a maximum-height jump and a standing long jump [136–138]. In these situations, the high-level task is relatively easy to specify. An explicit structure for the underlying controller can be assumed or a parameterized set of control signals can be used as design variables. Applying predictive simulation to locomotion is more challenging than tracking-based approaches and is also computationally expensive, but several recent studies show promise [89,139].

One proposed approach to understanding and simplifying neural control is the application of muscle synergies. In this growing area of research, muscle activity is decomposed using principal component analysis or related approaches to find groups of muscles that are activated together during movement, modulated by a single scaling parameter [140,141]. There is evidence that a consistent set of muscle synergies is used between tasks such as postural perturbations in a range of directions [142] or cycling at a range of speeds [143], and that the number/complexity of synergies is reduced in pathologies that affect the

motor control system, such as post-stroke [144]. Controversy remains, however, about whether synergies are truly representative of the underlying physiology or are merely a consequence of experimental data processing.

Assumptions and Limitations

Modeling neural control is a grand challenge. Fundamental questions about the nature of motor control remain unanswered, and as a result open questions remain about the form a model for neural control should take. How should we parameterize the neural control system in a model to capture the motions of interest with minimal complexity? What are the roles of reflexes and feedback? Can synergies help parameterize a neural control model? Initial results on the role of synergies and the presence/role of a central pattern generator (e.g., [145]) in driving motion show promise in helping to develop a simple model for the neural control of locomotion, but compelling evidence is still required to determine whether and how synergies should be incorporated into a model of physiological motor control. Much also remains to be learned about our internal "cost function" for various motion tasks (e.g., walking, running, and reaching). Outstanding questions include 1) how cost functions differ between individuals or with age and pathology; 2) whether the neural control system performs a real-time optimization when external or internal conditions change unexpectedly; and 3) if the human is optimizing in real time, what signals are the neural controller using to update and optimize the given cost function. More research is needed where experiments and simulations are combined for a wide range of motions to develop and validate better models of neural control.

3.5.2 Verifying Neural Control

Verification of neural control models is challenging because few laws or benchmark problems exist to characterize the underlying neural system. For tracking approaches, it is common to compare predicted muscle activations to a comparable optimal control solution to verify that coordination is similar when searching for local, rather than global, minimum effort (e.g., [99]). Investigators have also added noise to the input data to ensure tracking is robust [146]. For predictive simulation, the solution approach can be applied to simple models (e.g., a passive leg swing or simple maximal effort task) where the solution (i.e., the underlying control signal) is known. In the case of muscle synergies, verification should additionally confirm the eigendecomposition has been performed correctly (e.g., by reducing dimensionality, reconstituting the matrix, and checking the error relative to the original matrix) and that the error decreases as the number of synergy groups increases.

3.5.3 Validating and Evaluating the Robustness of Neural Control

Validation depends heavily on whether and how neural control is modeled in your study. In the case where experimental EMG data are used to estimate neural control, the same EMG data cannot be used for independent validation. In this case, *we recommend* comparing the resulting joint kinematics (if not prescribed) and net joint moments from the EMG-driven simulation to experimental kinematics and net joint moments from inverse dynamics (see Figure 6 for an example of joint moment comparison). Even with careful model calibration and EMG processing, it is generally unreasonable to expect that the EMG-driven moments will be an exact match to the inverse dynamics–based moments. For example, EMG is often

unavailable for small or deep muscles. In our group, we check for similar timing of moment peaks and strive for simulated EMG-driven moments within 2 standard deviations of inverse dynamics moments, unless known and documented modeling assumptions can account for additional differences (see the case study of Figure 6 for an example of this process).

In tracking simulations, muscle activations can be determined without using EMG data. In this case, *we recommend* comparing predicted muscle activity to measured EMG from your own experiments and/or previous studies. Since EMG is difficult to normalize and is subject to measurement error [147], you generally cannot compare EMG signals directly to predicted muscle activations from a simulation. Instead, you should determine whether onset/offset timing in experiments and simulations are in good agreement, after accounting for electromechanical delays between EMG and activation or force [148]. The case study of Figure 10 demonstrates the application of these principles to the study of running. Further, simulations should reproduce salient features of muscle coordination established in experimental studies (e.g., the vasti should show early stance phase activation in walking). If you test and simulate multiple conditions (e.g., walking with and without a carried load), your simulation should predict similar trends between EMG and simulated muscle activation (e.g., increased vasti activity [149]).

If simulation results do not match experimental EMG data, researchers will sometimes constrain muscles to be on or off at specified times in the simulation. In this case, the EMG data for these muscles become calibration data rather than validation data. Many tracking

simulations also apply extra forces or moments (e.g., reserve actuation in OpenSim terminology [6]) to account for the action of ignored passive structures or for insufficient muscle force-generating capacity. If your model includes reserve actuators, *we recommend* checking that these reserve moments are small relative to the total net joint moment. In our group, we aim for reserve joint torques (peak and RMS) smaller than 5% of net joint moments. In addition to comparing against EMG, *we recommend* comparing new predictions to existing simulations (i.e., validation against an independent implementation). Data from instrumented joint replacements, such as the Grand Challenge to Predict Knee Loads dataset [78], also provide an indirect set of independent validation data.

Predictive simulations allow for a larger set of independent validation data, since muscle activations and motions are typically resolved independently of EMG and motion capture data. *We recommend* making the following comparisons. First, onset and offset of muscle activity, as for tracking simulations, should be compared to experimental EMG. Kinematics, kinetics, and ground reaction forces should also be compared to experiments. Simulations should be within 2 standard deviations of experimental data if they are claimed to replicate human motion. In the case where a predictive simulation is designed to simulate a new motion that has not or cannot be measured, you should at least determine whether your framework is able to reproduce similar, known motions (e.g., a predictive simulation framework to design devices to aid human jumping should produce a human-like unassisted jump). As predictive simulation frameworks become more sophisticated and are able to synthesize a range of motions beyond just normal, unassisted walking, even more

data for independent validation will become available. For example, previous experimental studies have characterized how metabolic cost and muscle coordination change with speed and load. Researchers can now determine whether their predictive models show similar adaptations.

3.6 Numerical Methods for Simulation and Analysis

Often overlooked are the numerical methods used to solve for the unknowns in a simulation. For example, the multibody dynamics code generally forms a set of coupled differential-algebraic equations (DAEs), and numerical integration is required to find a solution (i.e., a trajectory of time-dependent states) that satisfies the system of DAEs. All numerical integrators are approximate, and their accuracy (how many digits of the solution are significant) or error tolerance (the absolute or relative difference with respect to the "true" answer) are parameters of the numerical integration algorithm. If the accuracy is too low, the solution may no longer be faithful to the modeled dynamics; if too high, the computation time will become prohibitive.

In musculoskeletal models, we are primarily concerned with inaccuracies in the model's state (e.g., joint angles and speeds, and muscle fiber lengths and activations). Errors in the state can result in erroneous forces, which lead to noisy acceleration estimates, poor numerical performance, and untrustworthy future states. For example, consider the case where one muscle in a model has a short optimal fiber length, say 5 mm. If the integration accuracy is 0.001, then each state variable is accurate to approximately three significant figures, or about 1 mm if lengths are stored in meters. In this case, an error of 0.001 m

amounts to 20% of the muscle's fiber length, which is sufficient to cause the muscle to switch from generating no passive forces to high passive forces, or from having a small force-generating capacity based on its force-length relationship to a capacity near its maximum. This behavior can cause the numerical solution to become unstable (i.e., chatter) and fail. *We recommend* choosing an accuracy corresponding to the largest change in a state variable that would result in an insignificant change in state derivatives.

Other commonly employed methods for determining unknowns in a model or simulation are root-solving and optimization algorithms. These methods are iterative and generally use derivatives of an objective function to steer the approach toward a solution. Optimization and root-solving methods are sensitive to several factors: derivative precision, function noise, and the convergence criteria used to terminate the search. Derivatives are often calculated numerically by perturbing the objective function, with smaller perturbations yielding higher-precision estimates of the derivative. However, if there is noise or imprecision in the evaluation of the objective function, then seemingly high-precision derivatives from these small perturbations will predominantly be a measure of the local noise and, thus, will be grossly inaccurate. For these algorithms to proceed efficiently and reliably, the perturbation size must be large enough to capture functional behavior and not merely noise. *We recommend* choosing a perturbation size that is substantially (e.g., an order of magnitude) larger than the known or estimated accuracy of the objective function [150]. A convergence criterion is used to assess how close the optimization or root-finding algorithm is to a final solution and specifies when the algorithm will terminate. *We recommend* verifying that the required convergence has been

achieved for any such computation by examining the plot of objective function value vs. iteration number.

3.7 Summary of Best Practices for Verification and Validation

In the sections above, we reviewed the current state of the art for selecting a modeling and simulation framework, and for verifying and validating its many possible outputs. The majority of modeling and simulation research falls into three categories: 1) building or adapting a musculoskeletal model for a particular application; 2) generating an inverse dynamic simulation based on measured experimental kinematics and/or kinetics; and 3) generating a forward dynamic simulation given a controller or estimated control signals. We, therefore, also provide in Tables 1, 2, and 3 a synopsis of the typical tests or outputs, dependencies, and validation best practices for each of these research categories. These tables can be used to guide the development of new models and simulations. The guidelines they contain will continue to evolve as we create new tools for modeling and simulation and new techniques and datasets for verification and validation.

4 Verification and Validation Challenges

The resources and best practices for verification and validation of models and simulations must continue to improve as the field progresses. We have identified several key challenges for verification and validation.

4.1 Creating "gold standard" datasets and benchmark problems

Gold standard datasets include comprehensive, high-fidelity experimental data collected using techniques inaccessible to many researchers (e.g., forces from instrumented joint replacements or motion from bone pin markers). These data should be made publicly available so others can use them to verify and validate both existing and new models. The Grand Challenge to Predict Knee Loads [78] is one example of a publicly available gold standard dataset. In addition to a comprehensive set of anthropometric and motion data, the Grand Challenge also includes a clear, validation-focused problem: predicting internal knee joint loads in a blinded manner. We need additional experimental datasets for validation and similar Grand Challenges to engage the field. For example, comprehensive datasets from animal models could provide full musculoskeletal geometry data and experimentally measured internal muscle and joint forces during motion. Comprehensive human datasets could include traditional motion capture data for a range of activities, along with internal muscle–tendon dynamics from ultrasound (e.g., [108]) and sarcomere lengths from recently developed optical microendoscopy techniques [151]. Accompanying these datasets should be a clear set of problems, such as predicting internal joint and muscle forces in the animal models or predicting changes in neural control for different activities (e.g., walking with and without an assistive device) in humans. Modeling and simulation will benefit from a clear set of benchmark problems to help establish baselines against which to compare both new and existing models and simulation tools.

4.2 Sharing models, data, and simulation tools

A related challenge for the field is further incentivizing and providing resources for sharing published models, data, and simulation tools with the wider research community. Sharing, although time intensive and not directly publishable, benefits the field by allowing others to reproduce, extend, and continue to verify and validate research products. Sharing can also benefit individual researchers by increasing awareness and subsequent citations of publications. Funding agencies have also increasingly pushed for open access to the results of sponsored research. We hope that the incentives and resources for sharing will continue to grow in the future, for example by encouraging open access as part of grant applications and journal publications. Contribution to and use of open source software should also be encouraged as a way of disseminating new tools that can immediately be used and extended by others.

4.3 Developing tools to help automate verification and validation

Verification tests are an integral part of current software packages for modeling and simulation. These test suites should continue to grow with the field, as new benchmark problems and gold standard datasets are made available and new modeling and simulation approaches are developed. The field also needs more tools to aid the model validation process, such as modern algorithms for efficient sensitivity testing (compatible with and/or embedded in NMS modeling packages) and automated tools to compare simulation results to established standards for kinematics, kinetics, and muscle activity and dynamics.

4.4 Learning and teaching others what's inside the "black box"

Knowledge of the physics and biology of the musculoskeletal system and its movement, as well as how existing tools transcribe real world systems into computational models, are the foundations for verification and validation. The verification and validation process demands an intimate knowledge of the capabilities and limitations of models and simulations. Members of the community must continue to expand their knowledge as new models, simulation techniques, and experimental modalities are developed. We also encourage the community to create and share teaching materials (e.g., hands-on tutorials and videos) for audiences from clinicians who are new to NMS modeling to advanced users and software developers.

4.5 Performing excellent research with modeling and simulation

Excellent research demonstrates the value of modeling and simulation by generating hypotheses and making predictions that impact real-world challenges. The set of available tools and datasets for building models, generating simulations, and validating predictions is rapidly growing. These advances are paving the way for modeling and simulation to be used as tools for planning musculoskeletal surgery, designing training programs to reduce sports injuries, and prototyping powered prosthesis that aid locomotion. As a field, we must continue to identify and solve pressing real-world challenges where modeling and simulation are used to uncover phenomena not accessible via observation or experiment alone.

5 Conclusions

Musculoskeletal models and dynamic simulations of movement provide powerful tools to study neuromuscular coordination, analyze athletic performance, and estimate internal loading of the musculoskeletal system. Simulations can also be used to identify the sources of pathological movement and establish a scientific basis for treatment planning. It is our responsibility as developers and users of these tools to ensure that our software has been thoroughly tested, our models and simulations accurately represent the essential physical phenomena, and that the conclusions we draw from these simulations are trustworthy. The guidelines provided here serve as a framework for achieving these goals.

6 Acknowledgments

We thank Michael Sherman, James Dunne, Edith Arnold, Samuel Hamner, and Bradley Davidson for their assistance with preparing this manuscript.

7 Funding

This work was supported by NIH grants U54 GM072970 and R24 HD065690, and the DARPA Warrior Web project.

8 Figure List

Figure 4: Elements of a musculoskeletal simulation. A model of the neuromusculoskeletal system can include computational models of muscle–tendon dynamics; geometry of bodies, joints, and muscles; models or estimates of contact; and models or estimates of neural control. A multibody dynamics engine is used to integrate the model's governing dynamic equations forward in time or solve for underlying motion and forces in an inverse analysis.

Figure 5: Case Study – Dynamic consistency and residuals

Figure 6: Case Study – Choosing and validating a musculoskeletal model

Figure 7: Verification test to ensure power from active and passive muscle fiber and tendon is equal to whole muscle actuator power. We generated a simulation with a constant muscle excitation of 0.6 (u), an initial block position of 0 m (x), and an initial block speed of 1 m/s (x_dot). We terminated the simulation after 0.5 s (t). The stacked area graph shows the summed power in the active muscle fiber (blue), passive muscle fiber (red), and tendon (green); the total muscle power is equal to the summation of these constituent powers (dashed black line).

Figure 8: Case Study – Tendon compliance sensitivity analysis

Figure 9: Case Study – Constraint-based contact modeling

Figure 10: Case Study – Comparing simulated muscle activations to EMG

9 References

[1] Nørgaard, B. L., Leipsic, J., Gaur, S., Seneviratne, S., Ko, B. S., Ito, H., Jensen, J. M., Mauri, L., De Bruyne, B., Bezerra, H., Osawa, K., Marwan, M., Naber, C., Erglis, A., Park, S.-J., Christiansen, E. H., Kaltoft, A., Lassen, J. F., Bøtker, H. E., and Achenbach, S., 2014, "Diagnostic performance of noninvasive fractional flow reserve derived from coronary computed tomography angiography in suspected coronary artery disease: The NXT trial (Analysis of coronary blood flow using CT angiography: Next steps).," J. Am. Coll. Cardiol., **63**(12), pp. 1145–55.

[2] Thacker, B. H., 2001, ASME Standards Committee on Verification and Validation in Computational Solid Mechanics.

[3] Anderson, A. E., Ellis, B. J., and Weiss, J. A., 2007, "Verification, validation and sensitivity studies in computational biomechanics.," Comput. Methods Biomech. Biomed. Engin., **10**(3), pp. 171–84.

[4] Henninger, H. B., Reese, S. P., Anderson, A. E., and Weiss, J. A., 2010, "Validation of computational models in biomechanics," Proc. Inst. Mech. Eng. [H], **224**(7), pp. 801–12.

[5] Lund, M. E., de Zee, M., Andersen, M. S., and Rasmussen, J., 2012, "On validation of multibody musculoskeletal models," Proc. Inst. Mech. Eng. [H], **226**(2), pp. 82–94.

[6] Delp, S. L., Anderson, F. C., Arnold, A. S., Loan, P., Habib, A., John, C. T., Guendelman, E., and Thelen, D. G., 2007, "OpenSim: open-source software to create and analyze dynamic simulations of movement.," IEEE Trans. Biomed. Eng., **54**(11), pp. 1940–50.

[7] Arnold, E. M., Ward, S. R., Lieber, R. L., and Delp, S. L., 2010, "A model of the lower limb for analysis of human movement.," Ann. Biomed. Eng., **38**(2), pp. 269–79.

[8] Arnold, E. M., Hamner, S. R., Seth, A., Millard, M., and Delp, S. L., 2013, "How muscle fiber lengths and velocities affect muscle force generation as humans walk and run at different speeds.," J. Exp. Biol., **216**(11), pp. 2150–60.

[9] Arnold, E. M., and Delp, S. L., 2011, "Fibre operating lengths of human lower limb muscles during walking.," Philos. Trans. R. Soc. Lond. B. Biol. Sci., **366**(1570), pp. 1530–9.

[10] Hamner, S. R., Seth, A., Steele, K. M., and Delp, S. L., 2013, "A rolling constraint reproduces ground reaction forces and moments in dynamic simulations of walking, running, and crouch gait.," J. Biomech., **46**(10), pp. 1772–6.

[11] Hamner, S. R., Seth, A., and Delp, S. L., 2010, "Muscle contributions to propulsion and support during running.," J. Biomech., **43**(14), pp. 2709–16.

[12] Hamner, S. R., and Delp, S. L., 2013, "Muscle contributions to fore-aft and vertical body mass center accelerations over a range of running speeds.," J. Biomech., **46**(4), pp. 780–7.

[13] Woltring, H. J., 1986, "A Fortran package for generalized, cross-validatory spline smoothing and differentiation," Adv. Eng. Softw. 1978, **8**(2), pp. 104–13.

[14] Wächter, A., and Biegler, L. T., 2005, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," Math. Program., **106**(1), pp. 25–57.

[15] Seth, A., Sherman, M., Reinbolt, J. A., and Delp, S. L., 2011, "OpenSim: a musculoskeletal modeling and simulation framework for in silico investigations and exchange," Procedia Iutam, **2**, pp. 212–232.
[16] Sherman, M. A., Seth, A., and Delp, S. L., 2011, "Simbody: multibody dynamics for biomedical research," Procedia Iutam, **2**, pp. 241–61.
[17] Reddy, M., 2011, API Design for C++, Elsevier, Burlington, MA, USA.
[18] Robinson, J. A., 2004, Software design for engineers and scientists, Elsevier, Jordan Hill, Oxford.
[19] Laz, P. J., and Browne, M., 2010, "A review of probabilistic analysis in orthopaedic biomechanics," Proc. Inst. Mech. Eng. [H], **224**(8), pp. 927–43.
[20] Hamby, D. M., 1994, "A review of techniques for parameter sensitivity analysis of environmental models.," Environ. Monit. Assess., **32**(2), pp. 135–54.
[21] Reinbolt, J. A., Schutte, J. F., Fregly, B. J., Koh, B. I., Haftka, R. T., George, A. D., and Mitchell, K. H., 2005, "Determination of patient-specific multi-joint kinematic models through two-level optimization.," J. Biomech., **38**(3), pp. 621–6.
[22] Valero-Cuevas, F. J., Johanson, M. E., and Towles, J. D., 2003, "Towards a realistic biomechanical model of the thumb: the choice of kinematic description may be more critical than the solution method or the variability/uncertainty of musculoskeletal parameters," J. Biomech., **36**(7), pp. 1019–30.
[23] Helton, J. C., and Davis, F. J., 2003, "Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems," Reliab. Eng. Syst. Saf., **81**(1), pp. 23–69.
[24] Box, G. E. P., and Tiao, G. C., 1992, Bayesian Inference in Statistical Analysis, Wiley.
[25] Haldar, A., and Mahadevan, S., 2000, Probability, Reliability, and Statistical Methods in Engineering Design, Wiley, New York.
[26] Melchers, R. E., 1987, Structural Reliability: Analysis and Prediction, Horwood, New York.
[27] Kepple, T. M., Arnold, A. S., Stanhope, S. J., and Siegel, K. L., 1994, "Assessment of a method to estimate muscle attachments from surface landmarks: A 3D computer graphics approach," J. Biomech., **27**(3), pp. 365–71.
[28] Cappozzo, A., Catani, F., Leardini, A., Benedetti, M. G., and Della Croce, U., 1996, "Position and orientation in space of bones during movement: experimental artefacts," Clin. Biomech., **11**(2), pp. 90–100.
[29] Lu, T. W., and O'Connor, J. J., 1999, "Bone position estimation from skin marker co-ordinates using global optimisation with joint constraints.," J. Biomech., **32**(2), pp. 129–34.
[30] Kane, T. R., Likins, P. W., and Levinson, D. A., 1983, Spacecraft Dynamics, McGraw Hill, New York.
[31] Schiehlen, W., ed., 1990, Multibody Systems Handbook, Springer-Verlag, Berlin.
[32] Hollars, M. G., Rosenthal, D. E., and Sherman, M. A., 1991, SD/FAST User's Manual, Symbolic Dynamics, Inc., Symbolic Dynamics, Inc., Mountain View, CA, USA.
[33] Seth, A., Sherman, M., Eastman, P., and Delp, S., 2010, "Minimal formulation of joint motion for biomechanisms.," Nonlinear Dyn., **62**(1), pp. 291–303.
[34] Shabana, A. A., 2013, Dynamics of Multibody Systems, Cambridge University Press, Cambridge.

[35] Chiari, L., Della Croce, U., Leardini, A., and Cappozzo, A., 2005, "Human movement analysis using stereophotogrammetry. Part 2: instrumental errors.," Gait Posture, **21**(2), pp. 197–211.
[36] Leardini, A., Chiari, L., Della Croce, U., and Cappozzo, A., 2005, "Human movement analysis using stereophotogrammetry. Part 3. Soft tissue artifact assessment and compensation.," Gait Posture, **21**(2), pp. 212–25.
[37] Pàmies-Vilà, R., Font-Llagunes, J. M., Cuadrado, J., and Alonso, F. J., 2012, "Analysis of different uncertainties in the inverse dynamic analysis of human gait," Mech. Mach. Theory, **58**, pp. 153–64.
[38] Silva, M. P. T., and Ambrósio, J. A. C., 2004, "Sensitivity of the results produced by the inverse dynamic analysis of a human stride to perturbed input data," Gait Posture, **19**(1), pp. 35–49.
[39] Schwartz, M. H., Rozumalski, A., and Trost, J. P., 2008, "The effect of walking speed on the gait of typically developing children.," J. Biomech., **41**(8), pp. 1639–50.
[40] Della Croce, U., Leardini, A., Chiari, L., and Cappozzo, A., 2005, "Human movement analysis using stereophotogrammetry. Part 4: assessment of anatomical landmark misplacement and its effects on joint kinematics.," Gait Posture, **21**(2), pp. 226–37.
[41] Kuo, A. D., 1998, "A least-squares estimation approach to improving the precision of inverse dynamics computations," J. Biomech. Eng., **120**(1), pp. 148–59.
[42] Thelen, D. G., Anderson, F. C., and Delp, S. L., 2003, "Generating dynamic simulations of movement using computed muscle control," J. Biomech., **36**(3), pp. 321–28.
[43] Crowninshield, R. D., and Brand, R. A., 1981, "A physiologically based criterion of muscle force prediction in locomotion," J. Biomech., **14**(11), pp. 793–801.
[44] McConville, J. T., Clauser, C. E., Churchill, T. D., Cuzzi, J., and Kaleps, I., 1980, Anthropometric Relationships of Body and Body Segment Moments of Inertia, Technical Report AFAMRL-TR-80-119, Anthropology Research Project, Inc., Yellow Springs, OH, USA.
[45] Zatsiorsky, V. M., and Seluyanov, V. N., 1983, "The mass and inertia characteristics of the main segments of the human body," Proceedings of the 8th International Congress of Biomechanics, H. Matsui, and K. Kobayashi, eds., Illinois, USA, pp. 1152–9.
[46] Chen, S.-C., Hsieh, H.-J., Lu, T.-W., and Tseng, C.-H., 2011, "A method for estimating subject-specific body segment inertial parameters in human movement analysis.," Gait Posture, **33**(4), pp. 695–700.
[47] Damavandi, M., Farahpour, N., and Allard, P., 2009, "Determination of body segment masses and centers of mass using a force plate method in individuals of different morphology.," Med. Eng. Phys., **31**(9), pp. 1187–94.
[48] Pataky, T. C., Zatsiorsky, V. M., and Challis, J. H., 2003, "A simple method to determine body segment masses in vivo: reliability, accuracy and sensitivity analysis," Clin. Biomech., **18**(4), pp. 364–8.
[49] Ehrig, R. M., Taylor, W. R., Duda, G. N., and Heller, M. O., 2007, "A survey of formal methods for determining functional joint axes.," J. Biomech., **40**(10), pp. 2150–7.
[50] MacWilliams, B. A., 2008, "A comparison of four functional methods to determine centers and axes of rotations.," Gait Posture, **28**(4), pp. 673–9.
[51] Lenaerts, G., Bartels, W., Gelaude, F., Mulier, M., Spaepen, A., Van der Perre, G., and Jonkers, I., 2009, "Subject-specific hip geometry and hip joint centre location affects calculated contact forces at the hip during gait.," J. Biomech., **42**(9), pp. 1246–51.

[52] Yamaguchi, G. T., and Zajac, F. E., 1989, "A planar model of the knee joint to characterize the knee extensor mechanism," J. Biomech., **22**(1), pp. 1–10.
[53] Steele, K. M., DeMers, M. S., Schwartz, M. H., and Delp, S. L., 2012, "Compressive tibiofemoral force during crouch gait.," Gait Posture, **35**(4), pp. 556–60.
[54] Siegler, S., Chen, J., and Schneck, C. D., 1988, "The three-dimensional kinematics and flexibility characteristics of the human ankle and subtalar joints—Part I: Kinematics," J. Biomech. Eng., **110**(4), pp. 364–73.
[55] Van Eijden, T. M. G. J., Kouwenhoven, E., Verburg, J., and Weijs, W. A., 1986, "A mathematical model of the patellofemoral joint," J. Biomech., **19**(3), pp. 219–29.
[56] Isman, R. E., and Inman, V. T., 1968, Anthropometric Studies of the Human Foot and Ankle, Technical Report 58, University of California, San Francisco and Berkeley.
[57] Lafortune, M. A., Cavanagh, P. R., Sommer, H. J., and Kalenak, A., 1992, "Three-dimensional kinematics of the human knee during walking," J. Biomech., **25**(4), pp. 347–57.
[58] Walker, P. S., Rovick, J. S., and Robertson, D. D., 1988, "The effects of knee brace hinge design and placement on joint mechanics," J. Biomech., **21**(11), pp. 965–74.
[59] Anderst, W., Zauel, R., Bishop, J., Demps, E., and Tashman, S., 2009, "Validation of three-dimensional model-based tibio-femoral tracking during running.," Med. Eng. Phys., **31**(1), pp. 10–6.
[60] An, K. N., Ueba, Y., Chao, E. Y., Cooney, W. P., and Linscheid, R. L., 1983, "Tendon excursion and moment arm of index finger muscles," J. Biomech., **16**(6), pp. 419–25.
[61] An, K. N., Takahashi, K., Harrigan, T. P., and Chao, E. Y., 1984, "Determination of muscle orientations and moment arms," J. Biomech. Eng., **106**(3), pp. 280–2.
[62] Spoor, C. W., and van Leeuwen, J. L., 1992, "Knee muscle moment arms from MRI and from tendon travel," J. Biomech., **25**(2), pp. 201–6.
[63] Arnold, A. S., Salinas, S., Asakawa, D. J., and Delp, S. L., 2000, "Accuracy of muscle moment arms estimated from MRI-based musculoskeletal models of the lower extremity.," Comput. Aided Surg., **5**(2), pp. 108–19.
[64] Delp, S. L., Loan, J. P., Hoy, M. G., Zajac, F. E., Topp, E. L., and Rosen, J. M., 1990, "An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures.," Ieee Trans. Biomed. Eng., **37**(8), pp. 757–67.
[65] Horsman, M. D. K., Koopman, H. F. J. M., van der Helm, F. C. T., Prosé, L. P., and Veeger, H. E. J., 2007, "Morphological muscle and joint parameters for musculoskeletal modelling of the lower extremity.," Clin. Biomech., **22**(2), pp. 239–47.
[66] Murray, W. M., Buchanan, T. S., and Delp, S. L., 2002, "Scaling of peak moment arms of elbow muscles with upper extremity bone dimensions," J. Biomech., **35**(1), pp. 19–26.
[67] Scheys, L., Van Campenhout, A., Spaepen, A., Suetens, P., and Jonkers, I., 2008, "Personalized MR-based musculoskeletal models compared to rescaled generic models in the presence of increased femoral anteversion: Effect on hip moment arm lengths.," Gait Posture, **28**(3), pp. 358–65.
[68] Holzbaur, K. R. S., Murray, W. M., and Delp, S. L., 2005, "A model of the upper extremity for simulating musculoskeletal surgery and analyzing neuromuscular control," Ann. Biomed. Eng., **33**(6), pp. 829–40.
[69] Sherman, M. A., Seth, A., and Delp, S. L., 2013, "What is a moment arm? Calculating muscle effectiveness in biomechanical models using generalized coordinates,"

Volume 7B: 9th International Conference on Multibody Systems, Nonlinear Dynamics, and Control, ASME, Portland, OR, USA, p. V07BT10A052.
[70] Jensen, R. K., 1989, "Changes in segment inertia proportions between 4 and 20 years," J. Biomech., **22**(6–7), pp. 529–36.
[71] Pavol, M. J., Owings, T. M., and Grabiner, M. D., 2002, "Body segment inertial parameter estimation for the general population of older adults," J. Biomech., **35**(5), pp. 707–12.
[72] Ganley, K. J., and Powers, C. M., 2004, "Determination of lower extremity anthropometric parameters using dual energy X-ray absorptiometry: the influence on net joint moments during gait," Clin. Biomech., **19**(1), pp. 50–6.
[73] Pearsall, D. J., and Costigan, P. A., 1999, "The effect of segment parameter error on gait analysis results," Gait Posture, **9**(3), pp. 173–183.
[74] Rao, G., Amarantini, D., Berton, E., and Favier, D., 2006, "Influence of body segments' parameters estimation models on inverse dynamics solutions during gait.," J. Biomech., **39**(8), pp. 1531–6.
[75] Grood, E. S., Suntay, W. J., Noyes, F. R., and Butler, D. L., 1984, "Biomechanics of the knee-extension exercise. Effect of cutting the anterior cruciate ligament," J. Bone Jt. Surg., **66**(5), pp. 725–34.
[76] Reinbolt, J. A., Haftka, R. T., Chmielewski, T. L., and Fregly, B. J., 2007, "Are patient-specific joint and inertial parameters necessary for accurate inverse dynamics analyses of gait?," Ieee Trans. Biomed. Eng., **54**(5), pp. 782–93.
[77] Bergmann, G., Graichen, F., and Rohlmann, A., 1993, "Hip joint loading during walking and running, measured in two patients," J. Biomech., **26**(8), pp. 969–90.
[78] Fregly, B. J., Besier, T. F., Lloyd, D. G., Delp, S. L., Banks, S. A., Pandy, M. G., and D'Lima, D. D., 2012, "Grand challenge competition to predict in vivo knee loads.," J. Orthop. Res., **30**(4), pp. 503–13.
[79] Kim, H. J., Fernandez, J. W., Akbarshahi, M., Walter, J. P., Fregly, B. J., and Pandy, M. G., 2009, "Evaluation of predicted knee-joint muscle forces during gait using an instrumented knee implant.," J. Orthop. Res., **27**(10), pp. 1326–31.
[80] Stansfield, B. W., Nicol, A. C., Paul, J. P., Kelly, I. G., Graichen, F., and Bergmann, G., 2003, "Direct comparison of calculated hip joint contact forces with those measured using instrumented implants. An evaluation of a three-dimensional mathematical model of the lower limb," J. Biomech., **36**(7), pp. 929–36.
[81] Herzog, W., and Read, L. J., 1993, "Lines of action and moment arms of the major force-carrying structures crossing the human knee joint.," J. Anat., **182**(2), pp. 213–30.
[82] Duda, G. N., Brand, D., Freitag, S., Lierse, W., and Schneider, E., 1996, "Variability of femoral muscle attachments," J. Biomech., **29**(9), pp. 1185–90.
[83] Raikova, R. T., and Prilutsky, B. I., 2001, "Sensitivity of predicted muscle forces to parameters of the optimization-based human leg model revealed by analytical and numerical analyses," J. Biomech., **34**(10), pp. 1243–55.
[84] Herzog, W., 1992, "Sensitivity of muscle force estimations to changes in muscle input parameters using nonlinear optimization approaches," J. Biomech. Eng., **114**(2), pp. 267–8.

[85] Ackland, D. C., Lin, Y.-C., and Pandy, M. G., 2012, "Sensitivity of model predictions of muscle function to changes in moment arms and muscle–tendon properties: A Monte-Carlo analysis.," J. Biomech., **45**(8), pp. 1463–71.
[86] Correa, T. A., Baker, R., Graham, H. K., and Pandy, M. G., 2011, "Accuracy of generic musculoskeletal models in predicting the functional roles of muscles in human gait.," J. Biomech., **44**(11), pp. 2096–105.
[87] Xiao, M., and Higginson, J. S., 2008, "Muscle function may depend on model selection in forward simulation of normal walking.," J. Biomech., **41**(15), pp. 3236–42.
[88] Cleather, D. I., and Bull, A. M. J., 2010, "Lower-extremity musculoskeletal geometry affects the calculation of patellofemoral forces in vertical jumping and weightlifting.," Proc. Inst. Mech. Eng. [H], **224**(9), pp. 1073–83.
[89] Wang, J. M., Hamner, S. R., Delp, S. L., and Koltun, V., 2012, "Optimizing locomotion controllers using biologically-based actuators and objectives," Acm Trans. Graph., **31**(4), pp. 1–11.
[90] Miller, R. H., Umberger, B. R., Hamill, J., and Caldwell, G. E., 2012, "Evaluation of the minimum energy hypothesis and other potential optimality criteria for human running.," Proc. R. Soc. B Biol. Sci., **279**(1733), pp. 1498–505.
[91] Handsfield, G. G., Meyer, C. H., Hart, J. M., Abel, M. F., and Blemker, S. S., 2014, "Relationships of 35 lower limb muscles to height and body mass quantified using MRI.," J. Biomech., **47**(3), pp. 631–8.
[92] Young, A., Stokes, M., and Crowe, M., 1984, "Size and strength of the quadriceps muscles of old and young women," Eur. J. Clin. Invest., **14**(4), pp. 282–7.
[93] Klein, C. S., Rice, C. L., and Marsh, G. D., 2001, "Normalized force, activation, and coactivation in the arm muscles of young and old men," J. Appl. Physiol., **91**(3), pp. 1341–49.
[94] Morse, C. I., Thom, J. M., Birch, K. M., and Narici, M. V, 2005, "Changes in triceps surae muscle architecture with sarcopenia.," Acta Physiol. Scand., **183**(3), pp. 291–8.
[95] Wickiewicz, T. L., Roy, R. R., Powell, P. L., and Edgerton, V. R., 1983, "Muscle architecture of the human lower limb.," Clin. Orthop., **179**, pp. 275–83.
[96] Ward, S. R., Eng, C. M., Smallwood, L. H., and Lieber, R. L., 2009, "Are current measurements of lower extremity muscle architecture accurate?," Clin. Orthop., **467**(4), pp. 1074–82.
[97] Zajac, F. E., 1989, "Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control," Crit. Rev. Biomedial Eng., **17**(4), pp. 359–410.
[98] Millard, M., Uchida, T., Seth, A., and Delp, S. L., 2013, "Flexing computational muscle: modeling and simulation of musculotendon dynamics.," J. Biomech. Eng., **135**(2), p. 021005.
[99] Anderson, F. C., and Pandy, M. G., 2001, "Static and dynamic optimization solutions for gait are practically equivalent," J. Biomech., **34**(2), pp. 153–61.
[100] Miller, R. H., Umberger, B. R., and Caldwell, G. E., 2012, "Limitations to maximum sprinting speed imposed by muscle mechanical properties.," J. Biomech., **45**(6), pp. 1092–7.
[101] Azizi, E., Brainerd, E. L., and Roberts, T. J., 2008, "Variable gearing in pennate muscles.," Proc. Natl. Acad. Sci. U. S. A., **105**(5), pp. 1745–50.
[102] Herzog, W., Lee, E. J., and Rassier, D. E., 2006, "Residual force enhancement in skeletal muscle.," J. Physiol., **574**(3), pp. 635–42.

[103] Rack, P. M. H., and Westbury, D. R., 1974, "The short range stiffness of active mammalian muscle and its effect on mechanical properties," J. Physiol., **240**(2), pp. 331–50.
[104] Rassier, D. E., MacIntosh, B. R., and Herzog, W., 1999, "Length dependence of active force production in skeletal muscle," J. Appl. Physiol., **86**(5), pp. 1445–57.
[105] Blemker, S. S., Asakawa, D. S., Gold, G. E., and Delp, S. L., 2007, "Image-based musculoskeletal modeling: applications, advances, and future opportunities.," J. Magn. Reson. Imaging, **25**(2), pp. 441–51.
[106] Blemker, S. S., and Delp, S. L., 2005, "Three-dimensional representation of complex muscle architectures and geometries," Ann. Biomed. Eng., **33**(5), pp. 661–73.
[107] Krylow, A. M., and Sandercock, T. G., 1997, "Dynamic force responses of muscle involving eccentric contraction," J. Biomech., **30**(1), pp. 27–33.
[108] Lichtwark, G. A., Bougoulias, K., and Wilson, A. M., 2007, "Muscle fascicle and series elastic element length changes along the length of the human gastrocnemius during walking and running.," J. Biomech., **40**(1), pp. 157–64.
[109] Scovil, C. Y., and Ronsky, J. L., 2006, "Sensitivity of a Hill-based muscle model to perturbations in model parameters," J. Biomech., **39**(11), pp. 2055–63.
[110] Brand, R. A., Pedersen, D. R., and Friederich, J. A., 1986, "The sensitivity of muscle force predictions to changes in physiologic cross-sectional area," J. Biomech., **19**(8), pp. 589–96.
[111] DeMers, M. S., Pal, S., and Delp, S. L., 2014, "Changes in tibiofemoral forces due to variations in muscle activity during walking.," J. Orthop. Res., **32**(6), pp. 769–76.
[112] Fukunaga, T., Roy, R. R., Shellock, F. G., Hodgson, J. A., and Edgerton, V. R., 1996, "Specific tension of human plantar flexors and dorsiflexors.," J. Appl. Physiol., **80**(1), pp. 158–65.
[113] Powell, P. L., Roy, R. R., Kanim, P., Bello, M. A., and Edgerton, V. R., 1984, "Predictability of skeletal muscle tension from architectural determinations in guinea pig hindlimbs.," J. Appl. Physiol., **57**(6), pp. 1715–21.
[114] Holzbaur, K. R. S., Delp, S. L., Gold, G. E., and Murray, W. M., 2007, "Moment-generating capacity of upper limb muscles in healthy adults.," J. Biomech., **40**(11), pp. 2442–9.
[115] Vidt, M. E., Daly, M., Miller, M. E., Davis, C. C., Marsh, A. P., and Saul, K. R., 2012, "Characterizing upper limb muscle volume and strength in older adults: A comparison with young adults.," J. Biomech., **45**(2), pp. 334–41.
[116] Correa, T. A., and Pandy, M. G., 2011, "A mass–length scaling law for modeling muscle strength in the lower limb.," J. Biomech., **44**(16), pp. 2782–9.
[117] Redl, C., Gfoehler, M., and Pandy, M. G., 2007, "Sensitivity of muscle force estimates to variations in muscle-tendon properties.," Hum. Mov. Sci., **26**(2), pp. 306–19.
[118] De Groote, F., Van Campen, A., Jonkers, I., and De Schutter, J., 2010, "Sensitivity of dynamic simulations of gait and dynamometer experiments to hill muscle model parameters of knee flexors and extensors.," J. Biomech., **43**(10), pp. 1876–83.
[119] Xiao, M., and Higginson, J., 2010, "Sensitivity of estimated muscle force in forward simulation of normal walking.," J. Appl. Biomech., **26**(2), pp. 142–9.
[120] Butler, D. L., Grood, E. S., Noyes, F. R., Zernicke, R. F., and Brackett, K., 1984, "Effects of structure and strain measurement technique on the material properties of young human tendons and fascia," J. Biomech., **17**(8), pp. 579–96.

[121] Maganaris, C. N., and Paul, J. P., 2002, "Tensile properties of the in vivo human gastrocnemius tendon," J. Biomech., **35**(12), pp. 1639–46.
[122] Maganaris, C. N., and Paul, J. P., 1999, "In vivo human tendon mechanical properties," J. Physiol., **521**(1), pp. 307–13.
[123] Bhargava, L. J., Pandy, M. G., and Anderson, F. C., 2004, "A phenomenological model for estimating metabolic energy consumption in muscle contraction," J. Biomech., **37**(1), pp. 81–8.
[124] Houdijk, H., Bobbert, M. F., and de Haan, A., 2006, "Evaluation of a Hill based muscle model for the energy cost and efficiency of muscular contraction.," J. Biomech., **39**(3), pp. 536–43.
[125] Lichtwark, G. A., and Wilson, A. M., 2007, "Is Achilles tendon compliance optimised for maximum muscle efficiency during locomotion?," J. Biomech., **40**(8), pp. 1768–75.
[126] Umberger, B. R., Gerritsen, K. G. M., and Martin, P. E., 2003, "A model of human muscle energy expenditure.," Comput. Methods Biomech. Biomed. Engin., **6**(2), pp. 99–111.
[127] Miller, R. H., 2014, "A comparison of muscle energy models for simulating human walking in three dimensions.," J. Biomech., **47**(6), pp. 1373–81.
[128] Lange, C., Martin, E., Piedbœuf, J.-C., and Kövecses, J., 2002, "Towards docking emulation using hardware-in-the-loop simulation with parallel platforms.," Proceedings of Fundamental Issues and Future Research Directions for Parallel Mechanisms and Manipulators, C.M. Gosselin, and I. Ebert-Uphoff, eds., Quebec City, Quebec, Canada, pp. 1–4.
[129] Fregly, B. J., Bei, Y., and Sylvester, M. E., 2003, "Experimental evaluation of an elastic foundation model to predict contact pressures in knee replacements," J. Biomech., **36**(11), pp. 1659–68.
[130] Neptune, R. R., and Hull, M. L., 1998, "Evaluation of performance criteria for simulation of submaximal steady-state cycling using a forward dynamic model," J. Biomech. Eng., **120**(3), pp. 334–41.
[131] Baraff, D., 1994, "Fast contact force computation for nonpenetrating rigid bodies," Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '94, ACM Press, New York, NY, USA, pp. 23–34.
[132] Lloyd, D. G., and Besier, T. F., 2003, "An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments in vivo," J. Biomech., **36**(6), pp. 765–76.
[133] Langenderfer, J., LaScalza, S., Mell, A., Carpenter, J. E., Kuhn, J. E., and Hughes, R. E., 2005, "An EMG-driven model of the upper extremity and estimation of long head biceps force.," Comput. Biol. Med., **35**(1), pp. 25–39.
[134] Ackermann, M., and van den Bogert, A. J., 2010, "Optimality principles for model-based prediction of human gait.," J. Biomech., **43**(6), pp. 1055–60.
[135] Erdemir, A., McLean, S., Herzog, W., and van den Bogert, A. J., 2007, "Model-based estimation of muscle forces exerted during movements.," Clin. Biomech., **22**(2), pp. 131–54.
[136] Van Werkhoven, H., and Piazza, S. J., 2013, "Computational model of maximal-height single-joint jumping predicts bouncing as an optimal strategy.," J. Biomech., **46**(6), pp. 1092–7.

[137] Anderson, F. C., and Pandy, M. G., 1999, "A dynamic optimization solution for vertical jumping in three dimensions.," Comput. Methods Biomech. Biomed. Engin., **2**(3), pp. 201–31.
[138] Ashby, B. M., and Delp, S. L., 2006, "Optimal control simulations reveal mechanisms by which arm movement improves standing long jump performance.," J. Biomech., **39**(9), pp. 1726–34.
[139] Geyer, H., and Herr, H., 2010, "A muscle-reflex model that encodes principles of legged mechanics produces human walking dynamics and muscle activities.," Ieee Trans. Neural Syst. Rehabil. Eng., **18**(3), pp. 263–73.
[140] Ting, L. H., and Macpherson, J. M., 2005, "A limited set of muscle synergies for force control during a postural task.," J. Neurophysiol., **93**(1), pp. 609–13.
[141] D'Avella, A., Saltiel, P., and Bizzi, E., 2003, "Combinations of muscle synergies in the construction of a natural motor behavior.," Nat. Neurosci., **6**(3), pp. 300–8.
[142] Torres-Oviedo, G., and Ting, L. H., 2010, "Subject-specific muscle synergies in human balance control are consistent across different biomechanical contexts.," J. Neurophysiol., **103**(6), pp. 3084–98.
[143] Wakeling, J. M., and Horn, T., 2009, "Neuromechanics of muscle synergies during cycling.," J. Neurophysiol., **101**(2), pp. 843–54.
[144] Cheung, V. C. K., Piron, L., Agostini, M., Silvoni, S., Turolla, A., and Bizzi, E., 2009, "Stability of muscle synergies for voluntary actions after cortical stroke in humans.," Proc. Natl. Acad. Sci. U. S. A., **106**(46), pp. 19563–8.
[145] Dzeladini, F., van den Kieboom, J., and Ijspeert, A., 2014, "The contribution of a central pattern generator in a reflex-based neuromuscular model," Front. Hum. Neurosci., **8**, p. 371.
[146] Seth, A., and Pandy, M. G., 2007, "A neuromusculoskeletal tracking method for estimating individual muscle forces in human movement," J. Biomech., **40**(2), pp. 356–66.
[147] Clancy, E. A., Morin, E. L., and Merletti, R., 2002, "Sampling, noise-reduction and amplitude estimation issues in surface electromyography," J. Electromyogr. Kinesiol., **12**(1), pp. 1–16.
[148] Corcos, D. M., Gottlieb, G. L., Latash, M. L., Almeida, G. L., and Agarwal, G. C., 1992, "Electromechanical delay: An experimental artifact.," J. Electromyogr. Kinesiol., **2**(2), pp. 59–68.
[149] Silder, A., Delp, S. L., and Besier, T., 2013, "Men and women adopt similar walking mechanics and muscle activation patterns during load carriage," J. Biomech., **46**(14), pp. 2522–2528.
[150] Gill, P. E., Murray, W., and Wright, M. H., 1982, Practical Optimization, Emerald Group Publishing, Bingley, United Kingdom.
[151] Llewellyn, M. E., Barretto, R. P., Delp, S. L., and Schnitzer, M. J., 2008, "Minimally invasive high-speed imaging of sarcomere contractile dynamics in mice and humans," Nature, **454**(7205), pp. 784–788.
[152] Farris, D. J., and Sawicki, G. S., 2012, "Human medial gastrocnemius force–velocity behavior shifts with locomotion speed and gait," Proc. Natl. Acad. Sci., **109**(3), pp. 977–982.

[153] Rubenson, J., Pires, N. J., Loi, H. O., Pinniger, G. J., and Shannon, D. G., 2012, "On the ascent: the soleus operating length is conserved to the ascending limb of the force–length curve across gait mechanics in humans," J. Exp. Biol., **215**(20), pp. 3539–3551.

Figure 1

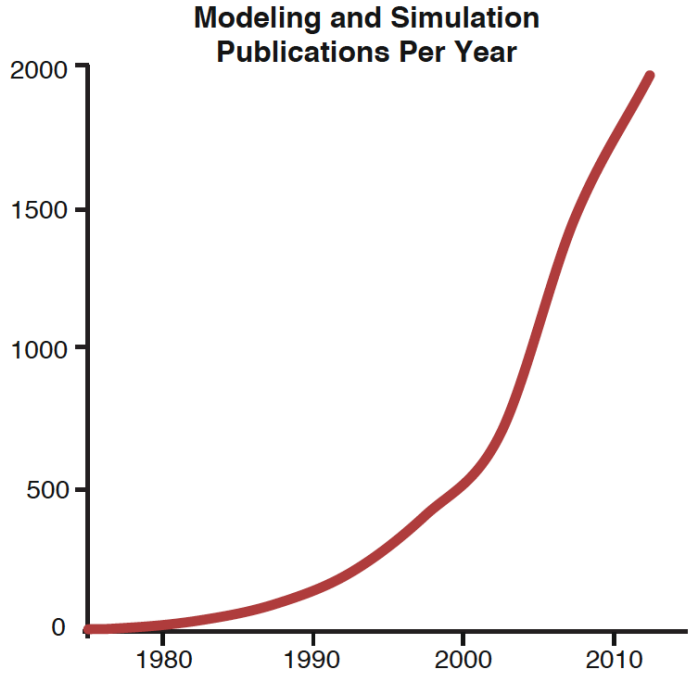

Modeling and Simulation
Publications Per Year
2000
1500
1000
500
0
1980
1990
2000
2010

Figure 2

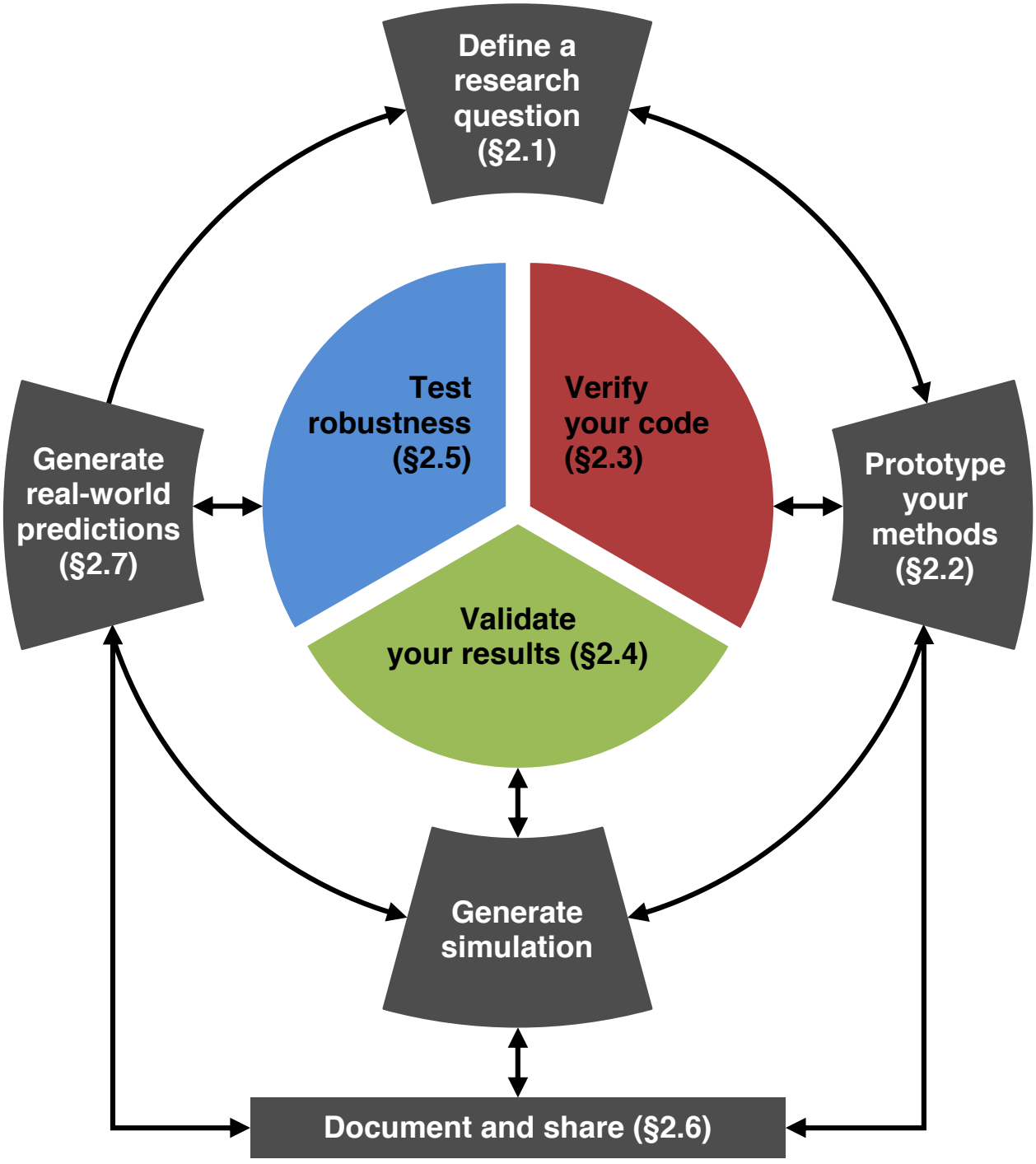

Define a research question (§2.1)
Test robustness (§2.5)
Verify your code (§2.3)
Generate real-world predictions (§2.7)
Prototype your methods (§2.2)
Validate your results (§2.4)
Generate simulation
Document and share (§2.6)

Figure 3

Introduction to verification and validation case studies

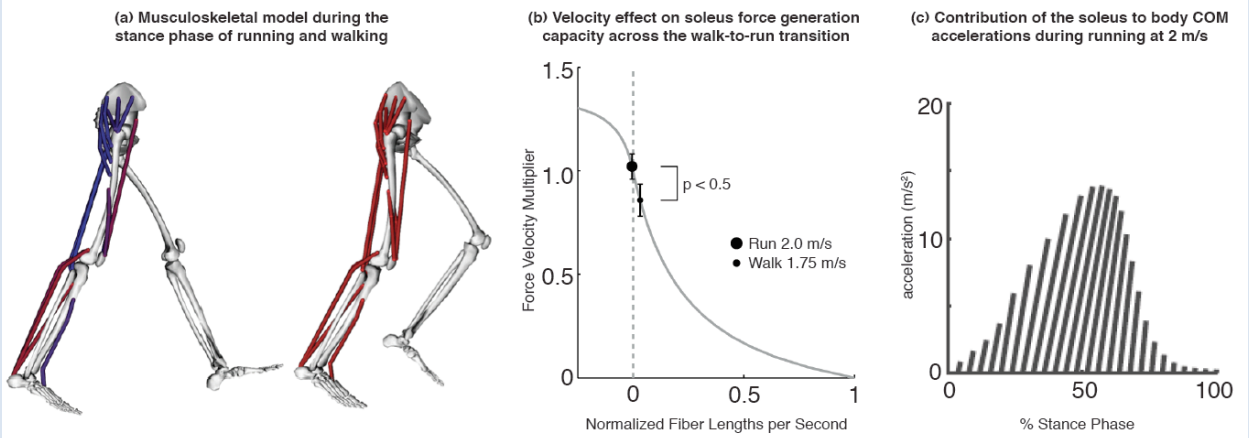
We will use two case studies to illustrate verification and validation best practices. These case studies provide example questions that are well suited to modeling and simulation. Consider the question, "Does switching from a fast walk to a slow run at a similar pace decrease muscle contraction velocities, enabling muscles to generate greater forces?" This question, tackled by Arnold and colleagues [8], had previously been unanswered—though experimental ultrasound data of plantarflexor muscle fascicles provided supporting evidence. This question is important for understanding human locomotion and also provided a valuable set of benchmark data describing muscle fiber velocities for both walking and running. Modeling and simulation were required to answer this question, since one cannot determine normalized velocities for a comprehensive set of lower extremity muscles without them. Finally, the question was testable using current modeling and simulation tools, though careful validation and sensitivity testing were necessary. Thus, this study satisfied each of the considerations described in Section 2.1.

This first research question, part of a series of investigations published by Arnold and colleagues [7-9], will serve as one of the case studies. The experimental protocol included collecting three-dimensional optical marker data, ground reaction forces, and electromyography (EMG) signals for 11 lower limb muscles during walking and running at a range of speeds, as shown in (a), including a fast walk and slow run at similar speeds. The outputs of the simulation were the fiber length and velocity of each muscle throughout the walking and running cycle, at each speed, which were used to answer the study's research question. The significant change in the force-generating capacity of the soleus muscle due to velocity effects across the walk-to-run transition is shown in (b).

A second research question suitable for modeling and simulation is, "Which lower extremity muscles brake, propel, and support the body mass center during running?" [11-12]. The answer to this question was unknown. Although examining joint powers and EMG provides some information, a model and simulation are required to quantify the complex skeletal and muscle–tendon dynamics during running, and to discover the contributions of individual muscles to center-of-mass (COM) accelerations. Hamner and colleagues were able to build on an existing modeling and simulation framework for analyzing accelerations induced by muscle forces, improving contact modeling to provide accurate answers for running at different speeds [10]. Finally, this study provided a valuable dataset for future analysis of running performance (e.g., to explore injury mechanisms and performance enhancement).

This research question is the basis for our second case study. The aim of this second investigation [11-12] was to determine the contribution of each lower extremity muscle to body weight support, COM braking, and COM propulsion. The investigators shared the experimental data used by Arnold and colleagues, focusing on the running trials. For this study, the outputs of the simulation were the forces generated by each lower extremity muscle and the induced COM acceleration generated by each of these muscles across the running cycle and across speeds. As an example, the upward and fore/aft COM accelerations generated by the soleus muscle are shown in (c).

See the pop-out boxes throughout the paper to learn more about the verification and validation used for each of these studies.



Adapted from [8,12].

Figure 4

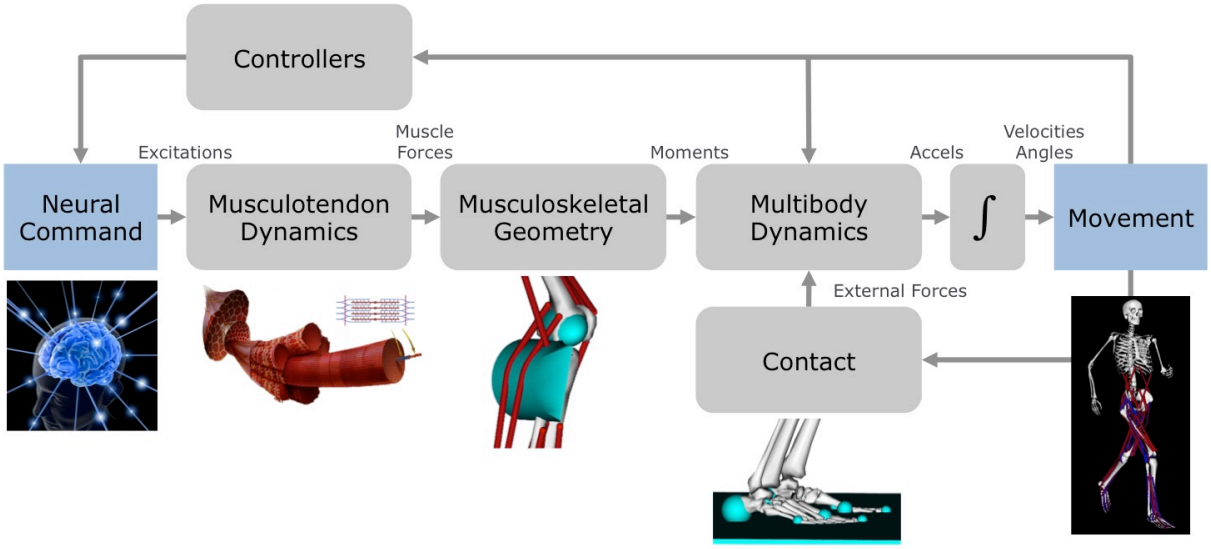

Controllers
Excitations
Muscle Forces
Moments
Accels
Velocities Angles
Neural Command
Musculotendon Dynamics
Musculoskeletal Geometry
Multibody Dynamics
∫
Movement
External Forces
Contact

Figure 5

Case Study – Dynamic consistency and residuals

In the investigations of running by Hamner and colleagues [11-12], and, indeed, in any study using motion capture data, a key validation step is ensuring the model, kinematics, and forces are dynamically consistent. This step is especially challenging when analyzing running, where the forces and accelerations are large. Before embarking on modeling and analysis, Hamner and colleagues ensured that the input data (marker trajectories and ground reaction forces) were collected using agreed-upon standards. Hamner scaled the model of each subject by precisely matching the virtual markers placed on the model to optical markers placed on the subject during the experiment. If the experimental data and model are poor representations of the measured motion and subject, producing an accurate simulation will not be possible.

Hamner and colleagues calculated the dynamic inconsistencies for their running simulations. Example force (a) and moment (b) inconsistencies from running at 4 m/s are shown in red. In OpenSim terminology, these dynamic inconsistencies are called “residual forces and moments” [6]. After careful model building and scaling, Hamner and colleagues performed a procedure to reduce residuals [6] by adjusting the body segment inertial parameters and kinematics slightly given estimates of marker measurement errors. The effect of these adjustments on the applied vertical force (a) and frontal plane moments (b) are shown in blue. By adjusting model parameters with significant uncertainty, such as the torso mass center location, and adjusting experimental kinematics within measurement error, Hamner and colleagues obtained models and kinematics that were dynamically consistent with measured ground reaction forces, reducing the effects of residual forces and moments on their simulation results. The torso mass center location was translated less than 5 mm and joint angles were changed by less than 1.5 degrees. Thus kinematics were still within measurement error, residual forces were much less than our recommended 5% of maximum external force magnitude (~100 N; dashed line in (a)), and residual moments were less than 1% of COM height times maximum external force magnitude (~35 Nm; dashed line in (b)).

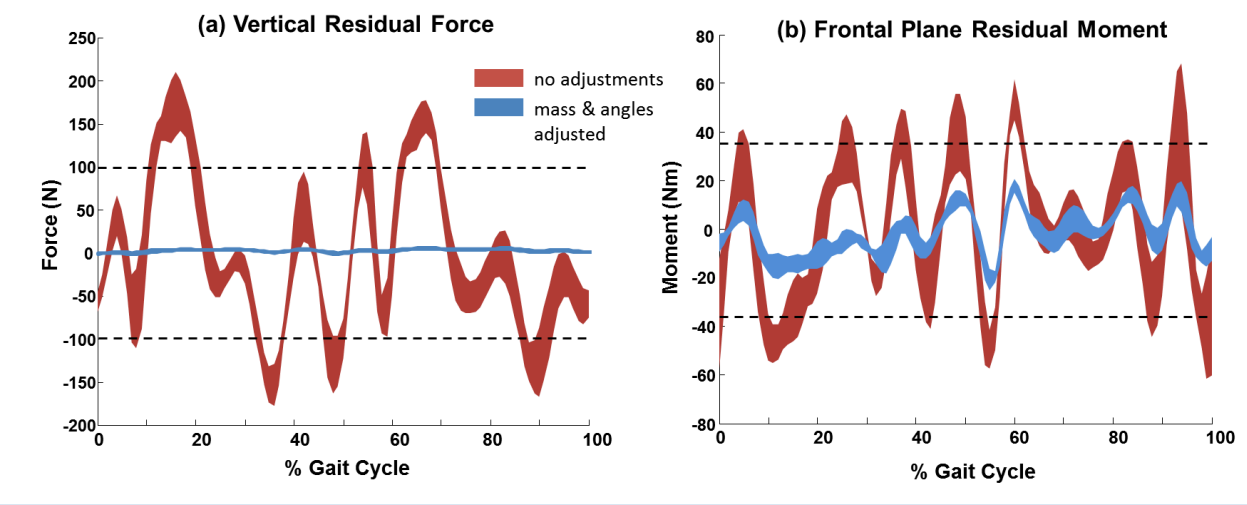
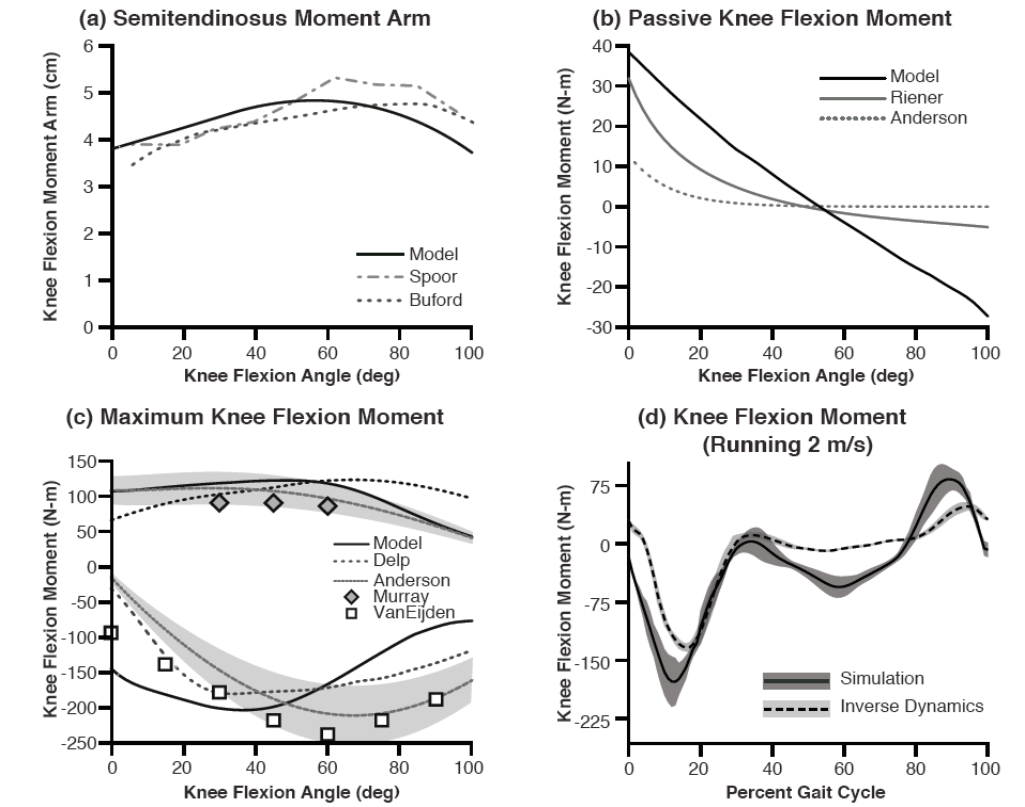


Figure 6

Case Study – Choosing and validating a musculoskeletal model

The investigation by Arnold and colleagues [8] into the force-generating capacity of lower extremity muscles depended on the interplay between muscle and tendon dynamics (i.e., fiber and tendon velocities) during motion. The researchers built a model [7] with high-fidelity musculoskeletal geometry and muscle–tendon dynamics based on a comprehensive muscle architecture dataset from twenty-one cadavers [95]. The investigators then validated the predictions of the model against a range of experimental data, including moment arm curves (a), passive moments (i.e., moments generated by muscles when they are inactive) (b), maximum isometric joint moments throughout the range of motion (c), and inverse dynamics moments from experimental studies of walking and running (d). The moment arms calculated with the model for the knee flexors (a), and other muscles, match experimental data measured in cadavers. The passive moments generated by inactive muscles crossing the knee (b) tend to generate larger forces than passive moments measured in experiments, especially at large knee flexion angles. This deviation occurs because the fibers of the vasti are stretched beyond their optimal lengths and generate passive forces. The maximum isometric moments generated by the knee flexors and extensors (c) have roughly the same magnitude as in experimental measurements, but the angle at which the moment peaks differs between experimentally measured maximum moments and the moments predicted by the model. These differences between the model and experimental results produce a discrepancy between the moments calculated from inverse dynamics and the moments generated in the simulation (d), and must be taken into account when interpreting the simulation results.



Adapted from [7, 8].

Figure 7

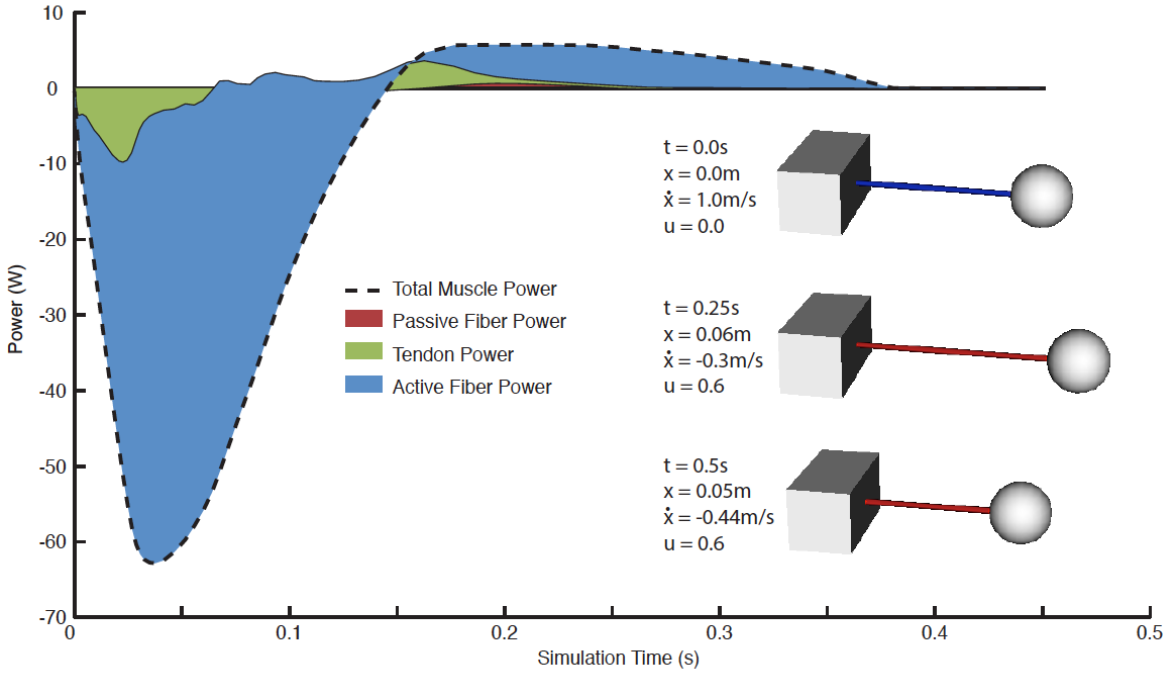
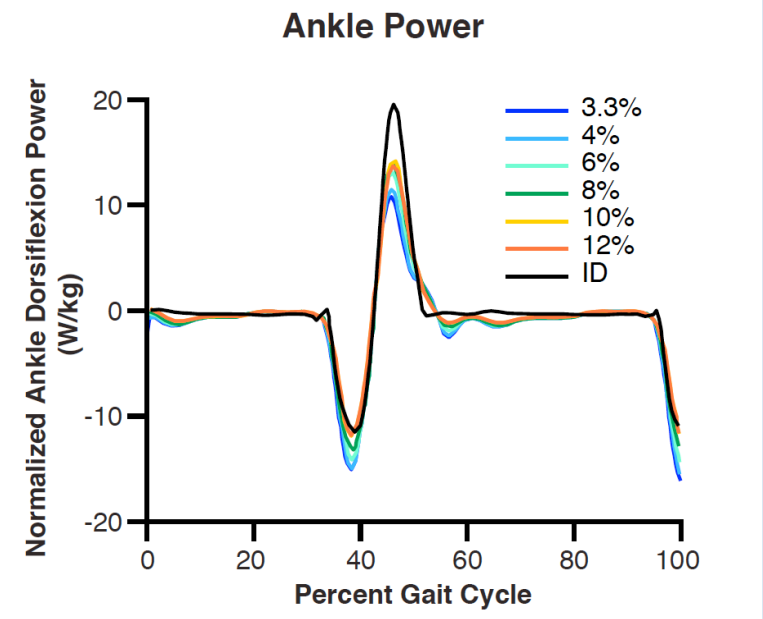

10
0
-10
-20
-30
-40
-50
-60
-70
Power (W)
0
0.1
0.2
0.3
0.4
0.5
Simulation Time (s)
Total Muscle Power
Passive Fiber Power
Tendon Power
Active Fiber Power
t = 0.0s
x = 0.0m
ẋ = 1.0m/s
u = 0.0
t = 0.25s
x = 0.06m
ẋ = -0.3m/s
u = 0.6
t = 0.5s
x = 0.05m
ẋ = -0.44m/s
u = 0.6

Figure 8

Case Study – Tendon compliance sensitivity analysis

Arnold and colleagues [8] selected a model whose musculoskeletal geometry and other muscle architecture parameters (e.g., optimal fiber lengths and pennation angles) were based on a high-quality dataset [95]. This choice was a first step in gaining confidence in the model's predictions of muscle–tendon dynamics. Since tendon properties are difficult to measure and can have an impact on simulation predictions, particularly for the plantarflexor muscles, the investigators performed a sensitivity analysis to determine the effect of tendon stiffness on plantarflexor muscle velocities and forces. They tested stiffness values for the plantarflexor muscles ranging from 3.3% (blue) to 12% (orange) tendon strain at maximum muscle isometric force, and examined the effect on predicted joint powers (normalized to body mass). A value of 10% gave total ankle power predictions that were the best fit to experimental data (i.e., net joint powers from inverse dynamics, shown in black) and tendon strains that were a good match to ultrasound measurements from previous experiments [152]. Arnold and colleagues also compared the operating lengths of the soleus to ultrasound measurements [153] and found that, in both simulation and experiments, the soleus operated on the ascending limb of the force–length curve for most of the time it was active during gait. Finally, changing the tendon stiffness of the plantarflexors did not change the study conclusions. Regardless of stiffness (within a reasonable range), the walk-to-run transition placed the plantarflexor muscles on a more favorable region of their force–velocity curve.

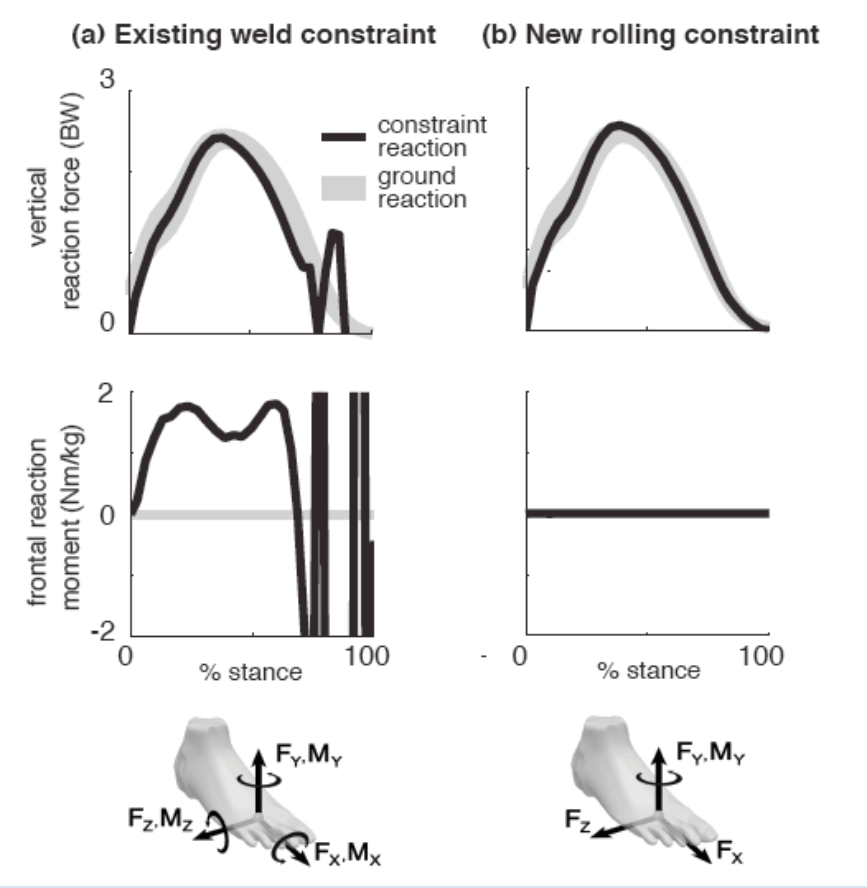


Adapted from [8].

Figure 9

Case Study – Constraint-based contact modeling

In the analysis of running by Hamner and colleagues [11,12], the investigation required calculating the contribution of each muscle to the acceleration of the center of mass of the body. When a muscle generates force, it induces a reaction force and moment at the ground; predicting these muscle-specific reaction forces and moments requires a foot–ground contact model, since the experimental ground reaction force measurement provides only the net values. In previous induced acceleration analyses for walking, it was assumed that the foot was welded to the ground during the stance phase of gait. Although reasonable for the single-limb stance phase of walking, assuming a weld constraint during running yielded unrealistic ground reaction force predictions. In particular, when Hamner and colleagues summed the contribution of each individual muscle to the ground reaction force, the resulting forces and moments were a poor match to the experimental net forces and moments (a). This result inspired Hamner and colleagues to implement a new foot–ground contact model (a rolling-without-slipping constraint), which yielded predictions much closer to experiments (b) [10]. Two of six potential reactions are presented: the vertical reaction force (first row) and the frontal plane reaction moment (second row) of the foot in contact with the ground. Model-generated reactions (thick black line) are compared to corresponding measured ground reactions averaged over 14 strides (shaded: mean ±1 standard deviation). Although the weld constraint matches most of the stance phase vertical ground reaction, it is evident that the weld applies a non-existent reaction moment in the frontal plane.

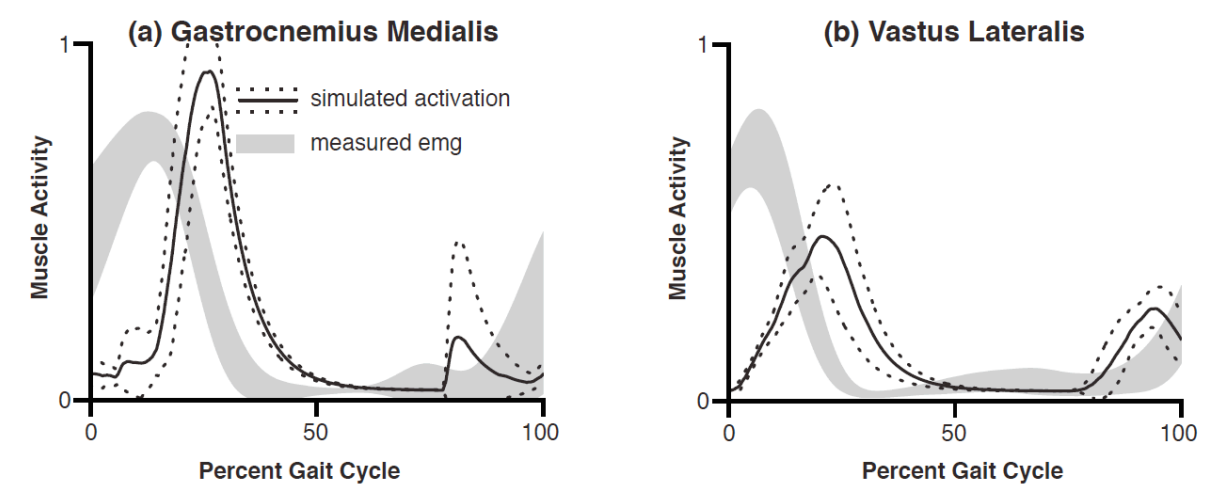


Adapted from [10].

Figure 10

Case Study – Comparing simulated muscle activations to EMG

In the running study of Hamner and colleagues [11,12], EMG data (gray shaded bands) were used to validate the model of neural control and predicted muscle activity (solid and dashed lines). In this study, the Computed Muscle Control tool included with OpenSim [42] was used to generate the set of muscle activations necessary to match the measured dynamic trajectory from each of the experimental running trials. The investigators compared the onset and offset timing of muscle activity, which aligned well with EMG once electromechanical delays (estimated to be around 75 ms [146]) were taken into account. They also determined that the simulations and experiments showed similar features, such as large activation of the plantarflexors (e.g., gastrocnemius medialis (a)) and quadriceps (e.g., vastus lateralis (b)) and increases in activation with running speed. Capturing these features of the muscle activations was important because these muscles made large contributions to the accelerations of the mass center; thus, large differences in the activations of these muscles would affect the conclusions of the study.



Adapted from [11].

Table 1: Summary of tests, dependencies, and validation best practices when building a musculoskeletal model or adapting a model to a new use. The first column lists tests to perform when building or adapting a musculoskeletal model. Each test should be performed throughout the range of motion for which the model will be analyzed. The second column lists the model parameters, experimental inputs, and antecedent output variables (the latter, in **bold**) being tested. Errors in each test can also stem from errors in each of the parameters and inputs listed in the second column (i.e., errors will propagate). Variables or inputs in *italics* are those to which an output is particularly sensitive. The third column lists best practices for validating the results of each test. These are recommended guidelines—if they cannot be achieved, the researcher should provide justification and perform sensitivity analyses to ensure that conclusions drawn from analysis of the model are robust given the errors or inaccuracies. The final column refers to sections of the article that provide more detail.

Musculoskeletal Model Test	Parameters and Inputs Tested	Validation Best Practices	Sec.
Model Kinematics: compute model kinematics through each joint's range of motion.	Joint definitions (e.g., joint location, orientation, and type); body segment lengths	Joint ranges of motion match experimental data. Modeled joints reproduce experimental motion from bone pin, cadaver, or imaging data to within measurement error.	3.1 3.2
Muscle Moment Arms: compute muscle moment arms throughout the model's range of motion.	Muscle geometry (attachment and via points, wrapping surfaces); **model kinematics**	Moment arms are within 2 standard deviations (SD) of experimental data measured by tendon excursion or MRI.	3.2
Forces from Muscle–Tendon Dynamics Model and constituent components, including **Active and Passive Muscle Fiber Force, Tendon Force and Strain, Muscle Fiber Lengths and Velocities:** use computational muscle–tendon dynamics model to replicate isolated muscle experiments.	Muscle–tendon dynamics model and its parameters, including maximum isometric force, pennation angle, *optimal fiber length*, passive muscle stiffness, force–length–velocity relationships, tendon stiffness, and *tendon slack length*	Difference between predicted and experimental muscle forces are within 10% of measured muscle force.	3.3
Passive Joint Moments: calculate the net moment generated by muscles and other modeled forces (e.g., ligaments) throughout joint range of motion when muscles have zero activation.	**Model kinematics**; **muscle moment arms**; muscle geometry; **muscle–tendon dynamics model** and its parameters, including passive force–length curve, *tendon slack length*, tendon stiffness, and maximum isometric force	Passive moment curves are within 2 SD of experimental data.	3.2
Maximum Net Joint Moments: calculate the net moment generated by muscles and other modeled forces (e.g., ligaments) throughout joint range of motion when agonist muscles have maximum activation input and all other muscles have zero activation.	**Model kinematics**; **muscle moment arms**; muscle geometry; **muscle–tendon dynamics model** and all related parameters	Maximum moment curves are within 2 SD of experimental data. Joint moments generated by the model during submaximal activation should also be tested (see Table 3).	3.2

Table 2: Summary of dependencies and validation best practices for common outputs from an inverse simulation, assuming the model under study has already been validated. The first column lists typical simulation outputs. The second column lists the model parameters, experimental inputs, and antecedent output variables (the latter, in **bold**) upon which each output depends. Errors in each output can also stem from errors in each of the dependencies listed in the second column (i.e., errors will propagate) and from errors in the model (see Table 1). Variables or inputs in *italics* are those to which an output is particularly sensitive. The third column lists best practices for validation. These are recommended guidelines—if they cannot be achieved, the researcher should provide justification and perform sensitivity analyses to ensure that conclusions are robust given the errors or inaccuracies. The final column refers to sections of the article that provide more detail.

Inverse Simulation Output	Dependencies	Validation Best Practices	Sec.
Kinematics of Joints, Bodies, and Points of Interest	*Joint definitions*; *body segment lengths*; sensor/motion capture marker placement; measurement noise; motion data processing	Difference between experimental and model sensors/markers is within measurement error. Kinematics are within 2 standard deviations (SD) of published data for similar motion.	3.2 3.1
Kinetics of Joints and Bodies, including net joint moments and residual forces and moments	**Kinematics**; *body segment inertial parameters*; *joint definitions*; measured forces and moments, including *points of application*, measurement noise, and data processing	Residual forces are < 5% of magnitude of net external force (peak and RMS). Residual moments are < 1% COM height times the magnitude of the net external force (peak and RMS). Net joint moments are within 2 SD of published data for similar motion.	3.1
Muscle–Tendon Unit Path Lengths and Lengthening Speeds	**Kinematics**; muscle geometry (attachment and via points, wrapping surfaces); joint definitions	Compare muscle paths and kinematics to imaging data (e.g., MRI).	3.2
Muscle Forces and/or Activations	*Tracking requirements* (e.g., **kinematics** and **joint moments**); optimization criterion (e.g., minimize sum of squared activations); muscle moment arms; muscle force-generating capacity as a function of **path lengths and speeds**; **joint definitions** (e.g., degrees of freedom and other passive structures modeled)	On and off timings of muscle activity are within electromechanical (EMG-to-force) delay (~100ms) of experimental EMG. Muscle activity and EMG curves are qualitatively similar. Net joint moments from muscles and other modeled structures account for 95% of tracked joint moments. Indirect: joint reaction force criteria are satisfied (see below).	3.3
Joint Reaction Forces	**Kinematics**; **kinetics**; ***muscle forces***; *joint definitions*; segment inertial properties; *muscle geometry*	Forces are within 2 SD of experimental joint forces (e.g., instrumented implants) for similar motion.	3.2 3.3 3.1.3

Table 3: Summary of dependencies and validation best practices for common outputs from a forward simulation, assuming the model under study has already been validated. The first column lists typical simulation outputs. The second column lists the model parameters, experimental inputs, and antecedent output variables (the latter, in **bold**) upon which each output depends. Errors in each output can also stem from errors in each of the dependencies in the second column (i.e., errors will propagate) and from errors in the model (see Table 1). Variables or inputs in *italics* are those to which an output is particularly sensitive. For a forward simulation, outputs additionally depend on the initial or current state of the model, which could include, for example, the coordinate positions and velocities and current muscle activations and fiber lengths. The third column lists best practices for validation. These are recommended guidelines—if they cannot be achieved, the researcher should provide justification and perform sensitivity analyses to ensure that conclusions are robust given the errors or inaccuracies. The final column refers to sections of the article that provide more detail.

Forward Simulation Output	Dependencies	Validation Best Practices	Sec.
Muscle Activation	Controller model or input controls (e.g., experimental EMG); activation dynamics model	On and off timings of muscle activation agree with experimental EMG. Muscle activation and EMG curves are qualitatively similar.	3.5
Muscle–Tendon Forces and their constituent components, including **Active and Passive Muscle Fiber Force, Tendon Force and Strain, Muscle Fiber Lengths and Velocities**	**Muscle activation**; initial/current kinematic state of the muscle; muscle–tendon dynamics model and its parameters, including maximum isometric force, pennation angle, *optimal fiber length*, passive muscle stiffness, force-length-velocity relationships, tendon stiffness, *tendon slack length*	Muscle fiber lengths and velocities and tendon strain are within 2 standard deviations (SD) of imaging data for similar motions. No unreasonably long or short muscle fibers or rapid changes in fiber lengths for motions such as walking.	3.3
Contact Forces and Moments	Initial/current model state and either *constraint type* (e.g., weld or rolling without slipping) or compliant contact model parameters (e.g., geometry, stiffness, dissipation, friction)	Contact forces and moments are within 2 SD of experimental data (e.g., ground contact forces and moments) for similar motion.	3.4
Kinetics of Joints and Bodies	**Muscle–tendon forces**; **contact forces**; initial/current model state; joint definitions; segment lengths and inertial parameters; muscle geometry	Joint moments are within 2 SD of published or independent experimental data for similar motion. If controller has a tracking objective, net joint moments from muscles and other modeled structures account for 95% of tracked joint moments.	3.1
Kinematics of Joints, Bodies, and Points of Interest	**Kinetics**; initial/current state; joint definitions; body segment lengths	Kinematics are within 2 SD of published or independent experimental data for similar motion. If controller has a tracking objective, difference between experimental and model sensors/markers is within measurement error.	3.1